

Gender and Number in Tunisian Arabic: A case of contextual
allosemy

Myriam Dali

A thesis submitted in fulfillment of the requirements
for the degree of Doctor in Philosophy in linguistics

Department of Linguistics

Faculty of Arts

University of Ottawa

© Myriam Dali, Ottawa, Canada, 2020

Abstract

In this thesis, I investigate the distribution and function(s) of gender in Tunisian Arabic. Based on the observation that gender can fill multiple functions and hence receive different interpretations in TA, I consider gender in this language to be subject to contextual allosemy. Allosemy is the equivalent of conditioned allomorphy at the level of LF (Wood, 2012; Marantz, 2013; Wood and Marantz, 2017) and is defined as a phenomenon in which a single morpheme can have multiple semantic realizations. My proposal is based on the observation that the interpretation of gender is conditioned by its syntactic environment, more specifically, by the class of the base noun and the function and interpretation of any node or abstract morpheme (e.g., number) that is parasitic of the n head hosting the class feature.

The current predominant view of gender in syntactic theory is that it has a nominalizing function (Lecarme, 2002; Kihm, 2005; Lowenstamm, 2008; Acquaviva, 2009; Kramer, 2009, 2014, 2015; Hammerly, 2018). According to this view, gender is hosted on the n head, which selects the root, and assigns it a nominal category (assuming DM). However, Arabic has other uses for the feminine gender, most of them associated with number. This, I argue, originates from the diachronic trajectory of the feminine affix *-a* in Proto-Semitic, where it started out as a nominalizer, then a singulative affix, then a group marker, before finally marking feminine gender (Hasselbach, 2014a,b; Dali and Mathieu, 2019a). These subsequent layers of meanings associated with what is now known as gender are all present in the synchronic picture of Arabic. Through an exhaustive inventory of data and diagnostics, I show that the role of gender is pervasive within the Tunisian Arabic DP, and is not limited

to nominalization. To account for these facts, I propose that gender is distributed among the different functional heads of the DP: Num, Q, D (see also [Farkas 1990](#); [Ritter 1993](#); [Giurgea 2008](#); [Croitor and Giurgea 2009](#); [Steriopolo and Wiltschko 2010](#); [Fassi Fehri 2012, 2018b,a](#); [Dali and Mathieu 2019a](#)).

The present thesis also focuses on plurals. I show that plural markers can also be distributed along the nominal spine ([Acquaviva, 2008](#); [Harbour, 2008](#); [Wiltschko, 2008, 2012](#); [Butler, 2012](#); [Mathieu, 2012, 2013, 2014](#); [Kramer, 2016](#)), accounting for their different functions (e.g., classifying, grouping, counting). These different functions associated with Arabic plurals are, I argue, due to the existence of a singulative operator that is not limited to the collective system, but is pervasive in Tunisian Arabic, as I show. Finally, these observations all suggest that one and the same abstract morpheme (e.g., the feminine and the plural) can receive different interpretations depending on the base they attach to and on their syntactic level, which motivates the allosemantic analysis put forward in this thesis.

ACKNOWLEDGEMENTS

I am immensely grateful to my wonderful supervisor, Éric Mathieu. Right from the beginning, he showed great faith in my potential to become a linguist. This gave me the confidence and the courage to carry on with my work throughout the years. He has been a truly empowering mentor and gave me invaluable advice and feedback that sharpened my writing skills and my scientific thinking. Thank you, Éric, for the enormous amounts of time and energy you invested in me, for your patience, your support and encouragement, your humour, and your generosity. I am grateful for our countless conversations about linguistics, literature, writing, travel, and life in general. You are a true role model for me, in academics and in life.

I am also deeply thankful to my committee members, Elizabeth Ritter, Ana Arregui, Andrés Salanova, and Dennis Ott. Prof. Ritter generously provided helpful feedback, which greatly contributed to shaping the thesis into its current form. Ana Arregui has been a great mentor and benevolent presence throughout my entire time at the University of Ottawa. She taught me to think “semantically” and she provided inestimable advice when supervising my second comprehensive exam. She has been extremely generous with her time, patiently witnessing my slow and gradual disentanglement of the Arabic number systems. Andrés Salanova was the first person I met when I first started my MA at the University of Ottawa. In the many classes that I had the chance to take with Andrés, I was struck by how he was always able to create a unique atmosphere in class, conducive to intellectual exchange and deep reflection. He was always available, with his door wide open and always had time for a chat. Andrés always encouraged me to think outside the box and I am very thankful

to him for that. I thank Dennis Ott for the great scientific rigour he instilled in me. I will always remember the great class he taught about the syntax of relative clauses, as he was starting out in our department. He helped me write a decent enough paper about a subject that was completely new to me, and his precious advice and critical thinking followed me throughout the subsequent years of my graduate studies.

It has been an honour for me to complete my graduate studies at the Linguistics department at the University of Ottawa. I have learned a lot from all faculty members. In particular, I wish to thank all those who taught me for the passion they instill into sharing their knowledge and experience on a daily basis: Stephen Levey, Marc Brunelle, Kevin McMullin, and Laura Sabourin. I was also lucky to get precious help from faculty members outside my department. I wish to thank Mireille Tremblay for being the first person who encouraged me to follow my passion for linguistics and for the tremendous help and support she has offered since my undergraduate years. I also thank Alan Bale, Paolo Acquaviva, and Rajesh Bhatt for insightful discussions who helped me shape my ideas.

I thank the administrative team at the Department of Linguistics and the Faculty of Arts: Monica Batallanos, Donna Desbiens, Maurice Bélanger, Roxanne Lacelle, and Paul-André David.

I'd like to gratefully acknowledge SSHRC and OGS's graduate scholarships, and various scholarships, travel grants, bursaries, and assistantships from the University of Ottawa for providing generous funding for my Master and my PhD. I also thank the Faculty of Graduate and Postgraduate studies for funding weekly writing cafes, intensive writing workshops, and writing retreats (via Thèsez-Vous, another incredi-

ble initiative!). These activities were extremely useful in breaking the isolation during the writing stage and in providing me with a good structure and a pleasant writing environment.

Thank you to my colleagues at the Linguistics department, for making these years highly memorable. In particular, I wish to thank my office mates Tharanga Weerasooriya, Gita Zareikar and Daiho Kitaoka for making our workspace a peaceful and pleasant one. Also, I thank Brandon Fry, Jumanah Abdulwahab, Basile Rousset, Myriam Lapierre, Ali Alamri, Emad Alansary, Félix Démeusles-Trudel, Mahsa Morid, Ray Therrien, and Yasmine Abou Taha, and Justin Case.

Salvio Digesto, you deserve a paragraph on your own. It is only when we became friends that I finally was able to call Ottawa home. I deeply cherish all these years spent together, as we encountered both success and failure along the way. Words cannot express my gratitude towards your sincere friendship.

My cousin Sarah, thank you for caring for me like a sister. You have always been a loving, caring presence in my life, and a pillar since I moved to Ottawa. Florian, you have always supported all the projects I have undertaken with great enthusiasm. Thank you for your constantly checking on me, making me laugh, and simply for being an amazing human being. I also thank my Ottawa friends: Aisha & Imran, Ivan & Anaïs, André, Leyda. Special thanks to my dear friend and neighbour Joanne Markle-Lamontagne who happened, not only to be a great linguist and talented writer, but also an incredibly generous friend who graciously offered to proofread my entire thesis and did it with great care and devotion. I also thank my friend Richard and my lovely Pélagie for all the joy you bring into our lives!

My Montreal friends: Amina is a reliable, caring and sincere friend, a sister, my awesome trio: Steph, Pav and Niku, thanks for all these years of fun, laughter, travels and great memories together. I am looking forward to more! Also thanks to Roberto, Martin, Marie-Noëlle, and Rémy. The Kanonkos, thank you for welcoming me as one of yours: Philippe, Carine, Billy, Taïna, and Russell. And of course, Pascale Audrey, Esmée and Caleb. Thanks to “les Amandines”: Amandine, Vanina, Floris, and Orphée.

To my Sivananda family, I am grateful for your presence in my life and to the inspiration and joy you bring. In particular, Atmaram, Karuna and Atmabala; Anjaneya, Yamuna, and Azzi, Lalita Chaitanya, Enoch, Swamiji, Gita, Françoise, Esha, Anshu, Romain, and Gopala.

My family back in Tunisia, there is not a day that goes by when I do not think of you. Thank you to the Zouaoui and the Dali families, from the bottom of my heart, for the unshakable love and support. In particular, I wish to honour the life and memory of two strong and wonderful women who left us during my PhD years: My aunt Bornia and my grandmother Wassila.

To my father, Ahmed, thank you for showing me unconditional love and giving me confidence in succeeding in all spheres of life. You are the funniest and most spontaneous person I know. Thank you for turning every one of my visits to Montreal into a big party. To my mother, Soumaya, thank you for being the most reliable person I know and for always listening to me and being genuinely happy when I succeed. My sister Yosra, thank you for being my first friend on this Earth, and for always being there for me. You hold the most special place in my heart and I wish

you all the success and joy in the world.

The last words go to my loving, talented, brilliant, beautiful life companion, Gino. Thank you for all the support and patience you showed during all these years. Life with you is rich and full of good surprises, and I cannot wait to see all the exciting things the future holds for us.

Contents

1	Introduction	1
1.1	Puzzles	1
1.2	The central problem and my proposal	5
1.3	Theoretical Assumptions	9
1.4	Tunisian Arabic	11
1.4.1	Overview and data sources	11
1.4.2	Root and pattern system	12
1.5	Overview of the thesis	13
2	Gender	15
2.1	Introduction	15
2.2	Gender assignment in TA	16
2.3	What is gender?	23
2.3.1	Defining gender	24
2.3.2	Gender as a nominalizer	31
2.3.3	Gender in root and pattern languages (Arad, 2005)	32

2.3.4	Class = n (Kihm, 2005)	39
2.3.5	Different flavours of n (Kramer, 2009, 2014, 2015)	41
2.4	Gender/number dependencies in TA	47
2.4.1	Certain assumptions about the structure of the DP	47
2.4.2	Gender is conditioned by number in TA inanimates	49
2.5	Proposal: uninterpretable gender on Num	56
2.5.1	Previous Num-based analyses of gender	57
2.5.2	A Num-based analysis of gender for TA inanimates	64
2.5.3	Alternative approaches	67
2.5.4	Summary	71
2.6	Conclusion	72
3	The singulative	74
3.1	Introduction	74
3.2	The collective system: definition and data	76
3.2.1	Singulative marking across languages	77
3.2.2	Definitions and properties of collective nouns	78
3.2.3	What the collective is not	88
3.2.4	The collective is a class type on n	96
3.2.5	Should we call it gender?	100
3.2.6	Arabic gender from a diachronic perspective	103
3.3	Proposal: interpretable gender on Num	105
3.3.1	The singulative is inflectional	106
3.3.2	The singulative: Interpretable gender on Num	109

3.3.3	Summary	114
3.4	Previous analyses	116
3.4.1	Greenberg (1972)	116
3.4.2	Ojeda (1992)	119
3.4.3	Zabbal (2002)	122
3.4.4	Borer and Ouwayda (2010)	126
3.4.5	Fassi Fehri (2004, 2012, 2018a,b)	129
3.4.6	Acquaviva (2008)	135
3.4.7	Mathieu (2012, 2013)	138
3.4.8	Kramer (2015)	142
3.5	Conclusion	144
4	Tunisian Arabic plurals	146
4.1	Introduction	146
4.2	Arabic plurals	147
4.2.1	Sound and broken plurals	148
4.2.2	Broken and sound plurals are equivalent	158
4.2.3	Summary	171
4.3	Plurals in a contrastive use	171
4.3.1	Data	172
4.3.2	Parallels with the collective system	178
4.3.3	Analysis: a reclassification of the noun	180
4.3.4	Summary	184
4.4	Broken plurals with a group reading and pluratives	184

4.4.1	Broken plurals with a group reading	185
4.4.2	Pluratives	189
4.4.3	Analysis: broken plurals are hybrid nouns	192
4.4.4	Summary	201
4.5	Conclusion	201
5	A Case of Allosemy	203
5.1	Introduction	203
5.2	What is Allosemy?	204
5.2.1	Potential pitfalls of an allosemic account	208
5.3	Allosemes in the DP	211
5.3.1	Saab and Lo Guercio (2019): Names as allosemes	211
5.3.2	The Num projection in Dagaare	213
5.4	Proposal: Allosemic morphemes in the TA DP	216
5.4.1	The feminine	216
5.4.2	The plural <i>-at</i>	220
5.4.3	Summary	223
5.5	Conclusion	224
6	Conclusion	226

List of Tables

2.1	Gender assignment on Tunisian Arabic sound plurals	22
2.2	Inflectional classes in Russian (from Corbett 2011)	27
2.3	The Arabic root /k t b/ and its patterns	33
2.4	Borrowed Hebrew nouns and verbs - Arad (2005, 35)	35
2.5	Tunisian Arabic borrowings from French.	36
2.6	Hebrew derivation through feminine	60
4.1	Tunisian Arabic canonical nouns and their plurals	174

Chapter 1

Introduction

1.1 Puzzles

This thesis presents an analysis of gender and number marking in Tunisian Arabic (henceforth, TA). I frame this dissertation around the puzzles presented below. All of these puzzles arise in the domain of the DP and are in some way connected to the coexistence of two number systems, each one using gender marking in different ways.

The first puzzle that I address has to do with gender marking on suffixal plurals of inanimate nouns. TA has two genders: masculine and feminine. The gender of animate nouns is assigned based on the sex of the referent. Inanimate nouns are assigned an arbitrary gender. While the gender of animate nouns remains unchanged across all numbers, all inanimate nouns are marked feminine in the plural, regardless of their gender in the singular. This is illustrated in (1)-(2), where both *babur* ‘boat’

The fifth puzzle that I address comes from the observation that it is possible for some nouns in TA to display both plural shapes (sound and broken) (7). Since it is traditionally assumed that each noun receives either one plural shape or the other, depending on phonological criteria, this data raises several issues. For instance, what conditions the choice of a plural shape over another, other than the phonology of the noun?

- (7) a. tfo*l* i-rabbi f-el **mfiiz**. [Tunisian Arabic]
 boy 3.MASC.SG-breed in-the goat.PL
 ‘A boy who breeds goats’
- b. ken fallaḥḥ ʕandu arð w ʃwayya **meʔz-et**.
 was farmer have-3.MASC.SG land and few goat-FEM.PL
 ‘He was a farmer who owned a land and a few goats.’

The puzzles presented above raise a number of questions that constitute the themes developed in this thesis. The main questions I address are the following:

1. What is the syntactic locus of gender in TA and how to account for its different interpretations?
2. How to formalize the close interactions between gender and number in the grammar?
3. How are the different noun classes organized in TA? What criteria determine the membership class of each noun?
4. What is the morphological nature of the broken (stem-internal) plural? Does it have special meaning? Does it have the same syntactic structure as the sound

(suffixal) plural?

Throughout this thesis, I address these questions by developing a heterogeneous analysis of gender and number marking in TA, arguing that the level of representation and interpretation is contingent on noun category. My hypotheses are presented in the next section.

1.2 The central problem and my proposal

An increased interest in the function of grammatical gender and other noun classification devices has emerged in recent years. A number of influential studies have looked at the origins and defining characteristics of gender (Dixon 1982, 1986; Corbett 1991; Aikhenvald 2000, among others). Others have focused on its function in the grammar (Picallo, 2008; Lecarme, 2002; Kihm, 2005; Lowenstamm, 2008; Acquaviva, 2009; Kramer, 2009, 2014, 2015; Hammerly, 2018). Within this area of investigation, many authors have questioned the semantic contribution that was traditionally assumed to be the hallmark of gender marking. Instead, they have pointed out gender's close interactions with nominalization (Lecarme, 2002; Kihm, 2005; Lowenstamm, 2008; Acquaviva, 2009; Kramer, 2009, 2014, 2015; Hammerly, 2018) and other functions related to the noun phrase (Farkas, 1990; Bernstein, 1993; Ritter, 1993; Fassi Fehri, 2018a). For example, it is now generally accepted that gender is a nominalizer which, assuming the Distributed Morphology approach (Halle and Marantz, 1993), is hosted on the n head, responsible for assigning a nominal category to the root.

linguists alike (Greenberg, 1972; Ojeda, 1992; Zabbal, 2002; Borer and Ouwayda, 2010; Fassi Fehri, 2012, 2018a,b; Mathieu, 2012, 2014). However, to date, no study has specifically looked at the function of gender in these systems, and at the diverse issues this marking system raises both on theoretical and empirical perspectives. This dissertation aims to fill this gap.

The central claim of this dissertation is that gender in TA cannot be reduced to a unique function (e.g., nominalization). Rather, the role of gender is contingent on the category of noun to which it is attached. To support this claim, I rely on the notion of *contextual allosemy*, where a single morpheme can have multiple semantic realizations (the parallel of contextual allomorphy in the LF domain). Put simply, I propose that when attached to a noun belonging to a certain class (i.e., the count class), gender has a nominalizing function that can be endowed with a semantic content, depending on whether the noun is animate or not. In the context of the other class (the collective class), gender has a different, inflectional function, and biological sex distinctions are neutralized. The notion of allosemy constitutes the main thread of this thesis, bringing the different chapters together. Other claims emanate from this central proposal.

First, the nature of nouns constituting the collective class is scrutinized, based on assumptions about their characteristics (i.e., collective nouns are notoriously claimed to form a homogeneous semantic class). I propose that, although there seems to be some semantic coherence within the collective class, the latter is better defined as a morphological class, with the essential characteristic of marking its individual-denoting nouns with the feminine suffix.

Another general theme in this dissertation concerns the level of representation for each instance of gender and number marker. For example, some plurals are considered to be lexical (Lecarme, 2002; Acquaviva, 2008; Lahrouchi and Lampitelli, 2014; Lahrouchi and Ridouane, 2016) and are therefore claimed to be hosted on n . Gender, which is generally viewed as a derivational morpheme, is also considered to be hosted on n . Therefore, many instances of morphemes involving information about both gender and number are often assumed to be derivational in nature (Lecarme, 2002; Kihm, 2005; Acquaviva, 2008, 2009; Kramer, 2009, 2014, 2015). I argue, however, that both gender and number have instances that can be distributed among different levels of the nominal spine, based on the base they attach to and the resulting function. For example, gender can be encoded on n or Num, depending on its function in that specific instance. Similarly, the plural can be encoded on n , Num or Q, based on its interpretation and the productivity of its shape and meaning (see also Acquaviva 2008; Harbour 2014; Wiltschko 2008, 2012; Kramer 2016 for distributed plural proposals).

To sum up, this dissertation aims to sort out the intricate gender and number systems of TA, and most importantly, to formalize them in a way that reduces all their complexities to the simple encoding of certain sets of features on the right functional heads. In the end, the learning of this complex system boils down to accurately encoding the grammatical classification of each noun on n . This information, coupled with world knowledge (i.e., whether the noun refers to an animate entity), is enough to derive the right function and interpretation for small pieces of conditioned meaning like gender and number morphemes. Speakers can compute meaning out of

a combination of noun class and gender/number morpheme thanks to their linguistic knowledge about their native language.

1.3 Theoretical Assumptions

Throughout this thesis, I adopt the assumptions of the Minimalist Program (Chomsky, 1995, 2000, 2001), supplemented with the architecture of Distributed Morphology (DM) (Halle and Marantz, 1993).

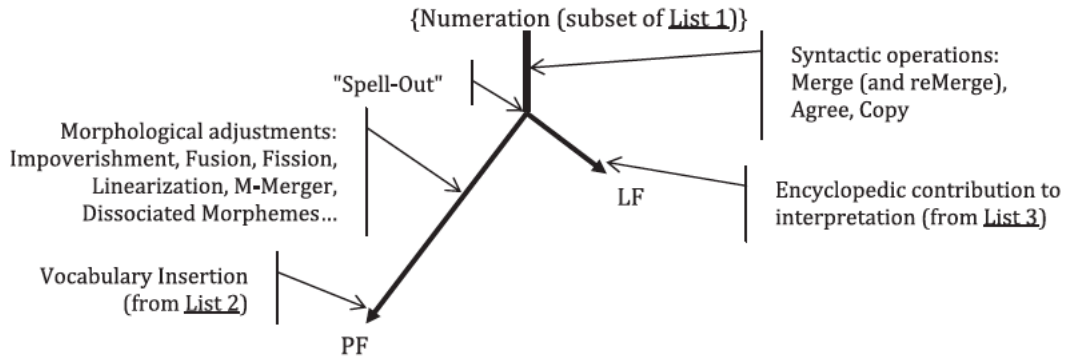
The Minimalist Program aims to achieve a minimal formulation of grammatical theory by reducing the set of theoretical mechanisms to a bare minimum. Chomsky proposes that the language faculty involves a computational system that feeds into the two components of the mind dealing with sound and meaning: the articulatory-perceptual (A-P) system and the conceptual-intentional (C-I) system. The computational system of human language interacts with these systems through two distinct interface levels: Phonetic Form (PF) and Logical Form (LF). The sound-meaning link is described as a derivation, taking a single array of lexical elements as its output. The two interface representations are different and one is not derived from the other. Syntactic structures are interpreted semantically at LF, and are assigned phonological material at PF. The point in the derivation where computation splits is called “Spell Out”.

In the derivation, the relations between the elements in the Numeration (the lexical items that will be used in the syntactic derivation) are made explicit by combining the various elements up in a phrase representation. This is done through the

basic compositional operation MERGE. Another operation, AGREE, relates features of syntactic objects (Chomsky, 2000, 2001). In Agree, a probe (a syntactic head with unvalued ϕ -features) establishes an Agree relation with a head carrying the inherent ϕ -features, called the goal, in order to get these features.

By adopting DM, I assume that words are also built in the syntactic component, following the Y-model described below, where syntax feeds into two distinct interface levels (10).

(10) The model: DM by Halle and Marantz (1993) (from Harley 2014, 228)



According to DM, the syntactic component can only manipulate terminal nodes that consist of formal features (or bundles thereof), which do not have phonological material when they enter the derivation. Once the syntactic operations are completed, morphological operations can manipulate feature bundles at the terminal nodes in a post-syntactic component. After the completion of all syntactic and post-syntactic operations, vocabulary items, which have a phonological form, are inserted into the terminal nodes in the Spell-out process. Vocabulary Insertion (VI) follows

Halle’s (1997) Subset Principle, which specifies that the phonological exponent of a vocabulary item can be inserted if the item contains all or a subset of the features present at the terminal node. At the same time, the item must have no feature that is absent from the node. In the case where several items compete for insertion, the one that matches the most features of the terminal node will be inserted. A single morpheme can have different alternative realizations depending on the phonological or morphological context in which it appears, or even the presence of another morpheme. The latter concept is of particular importance to the notion of allosemy, which I discuss in [Chapter 5](#). Allosemy is the LF counterpart of allomorphy, where a single morpheme can give rise to multiple semantic interpretations. One crucial assumption in DM is that lexical categories are composed of a categorizing head and a category-neutral root (Marantz 1997, 2001; Arad 2003, 2005; Embick and Noyer 2007; Harley 2014, among others). For example, nouns are built by merging a nominalizing head (*n*) with a category-neutral root. This assumption is essential to the topics investigated in this dissertation and plays a crucial role in [Chapter 2](#), where it is discussed in more detail.

1.4 Tunisian Arabic

1.4.1 Overview and data sources

Tunisian Arabic is a modern dialect spoken by over 11 million speakers in Tunisia, North Africa. TA is mostly intelligible to speakers of other Maghrebi dialects (e.g., dialects spoken in Morocco, Algeria, Libya, Mauritania). TA is an SVO and (most of

the time) a null-subject language. The dialect is characterized by frequent borrowings from French and Italian, due to the country's past history as a French protectorate and its geographic contiguity to Italy. Most of the phenomena (i.e., collective nouns, gender assignment, plural assignment, etc.) discussed in this dissertation is also characteristic of Standard Arabic and modern dialects. The object of study is the Tunisian dialect, due to the availability of data and judgments, but most claims are valid for all other dialectic variants.

The data originates from three main sources: my own judgments as a native speaker, the output of other native consultants, and the Online Tunisian Arabic Corpus (<http://www.tunisiya.org/>). The latter is the fruit of a project led by Karen McNeil and Miled Faiza, who compiled no less than 2,006 texts, comprising 881,964 words from multiple sources, such as phone conversations, TV and radio shows, blogs, recipes, folktales, etc.

1.4.2 Root and pattern system

Arabic is a root and pattern system, which is characteristic of Semitic languages. In root and pattern languages, the root consists of a set of consonants arranged in a specific sequence. This root identifies the general realm/concept of the word's meaning. Additional information (typically grammatical, such as the category or tense) is represented in the stem's vocalic and syllabic features, called "the pattern". Stems in Arabic and other Semitic languages indicate different grammatical contexts by using the root and pattern system and as a result can appear in quite different shapes. Compare, for example, the many variants of the famous Arabic root */k-t-*

b/: the past stem (active) *katab-*, as in *katab-tu* ‘I wrote’; the past stem (passive) *kutub-*, as in *kutib-a* ‘it was written’; the present stem (active) *-ktub-*, as in *a-ktub-u* ‘I write’, and the active participial stem *kaatib-*, as in *kaatib-un* ‘writing [one]’.

The study of root and pattern morphology easily lends itself to illustrating some DM claims due to the phonological distinguishability of roots from actual “words”. The assumption that roots are devoid of grammatical categories is both phonologically and semantically noticeable, since roots in Semitic are unpronounceable and lack a fixed meaning prior to their combination with a syllabic pattern. Semitic roots also lack a category. Noun and verb-creating morphology is found in vocalic patterns and is clearly distinguished from the root, which also plays in favour of a DM approach for Arabic (See [Arad 2005](#) for a DM study of root and pattern morphology, with a focus on Hebrew). These facts are of particular importance when it comes to investigating notions such as the locus of gender, since the latter is also noticeably absent from the root.

1.5 Overview of the thesis

The structure of this thesis is as follows. In [Chapter 2](#), the notion of grammatical gender is defined, and I discuss its function in the grammar of Arabic. I also address the interaction between gender and number in TA through the first puzzle, which is raised by the fact that all inanimate sound plurals take the feminine shape. I conclude this chapter by proposing that gender features can be found on the Num head.

In [Chapter 3](#), the nature of collective nouns is discussed. I explore the possible criteria for noun classification in TA, arguing that this categorization is not a semantic, but rather a morphological one. [Chapter 3](#) also addresses the locus of representation and interpretation of the feminine in nouns of the collective class. I argue that gender is interpretable on Num in collective nouns.

[Chapter 4](#) presents the case of broken plurals in TA and discuss their morphological nature, by asking the question: are broken plurals derivational or inflectional? This chapter also explores puzzles relative to agreement triggered by broken plurals in TA. It is argued that broken plurals nouns are underlyingly feminine, based on diachronic facts. I discuss another puzzle pertaining to the fact that some TA nouns can alternate between two plural shapes and argue that this alternation is conditioned by semantic factors.

In [Chapter 5](#), I formalize the generalizations stated throughout the thesis, that noun category determines the interpretation gender in TA. For this purpose, I resort to the notion of allosemy, whereby one single morpheme can be associated to several semantic readings, depending on its morphological environment. This chapter aims to formalize the generalization stated above, namely, that noun category determines the interpretation of gender in TA. Focusing on the feminine gender, I argue that the different meanings associated with the [+FEM] feature are allosemic.

Finally, in [Chapter 6](#), I summarize and discuss the conclusions of this research and future lines of research.

Chapter 2

Gender

2.1 Introduction

In this chapter, I examine the gender assignment system in TA nominals. This system is complex, because Arabic has two parallel nominal systems, each one using and interpreting gender in different ways. The coexistence of two systems using the same feature for two distinct ends raises interesting questions, namely about how a language speaker systematically distinguishes between nouns of the different systems and, consequently how they interpret gender features according to the right system. This chapter examines the first system, referred to here as the *count system*, where gender can be natural or arbitrary. In this class, interesting interactions are found between gender and number, involving animacy and interpretability. This phenomenon is at the core of the discussion about gender in the TA count system, since it raises questions about the definition of gender, its role, and its locus in the

DP.

In [Section 2.2](#), I present the main descriptive facts about gender in TA. [Section 2.3](#) looks at previous definitions of gender and other noun classification devices. I also argue that gender has a nominalizing function and is consequently hosted on *n*, the nominalizing head. In [Section 2.4](#), I address the case of gender-number interactions in TA and other languages, and propose that these dependencies result from the presence of gender features on the Num head. [Section 2.5](#) concludes with a summary of the main claims about gender in TA and its syntactic locus.

2.2 Gender assignment in TA

This section is about gender assignment in the “count system”, in contrast with the “collective system”. These appellations are simplified ones that refer to two parallel number systems that coexist in Arabic. [Chapter 3](#) discusses the collective system, and presents a new dimension of gender assignment in Arabic. I also discuss the nature of the different noun classes in Arabic in more detail. In the count system, number is marked by suffixes/stem shapes dedicated to number. In the collective system, all nouns are singular (both morphologically and syntactically) and gender marking is used to distinguish between individual and sum reference. Here, I present the main facts about gender assignment in the count system in order to further propose an analysis that accounts for the different functions and interpretations of gender in TA.

Arabic distinguishes two genders: masculine and feminine. For animate nouns,

gender is assigned according to the biological sex of the referent. In mating pairs (Harris, 1991), the masculine form is unmarked, while the feminine form is marked by the ending *-a* (1). ¹

- (1) *Mating pairs: masculine feminine*
- | | | | |
|--------|----------------------|----------|------------------------|
| qerd | ‘(male) monkey’ | qerd-a | ‘(female) monkey’ |
| xal | ‘uncle’ | xal-a | ‘aunt’ |
| ħajjem | ‘(male) hairdresser’ | ħajjem-a | ‘(female) hairdresser’ |

Some male and female pairs have suppletive forms, where each gender is associated to a different lexical item. (2)

- (2) *Suppletive pairs: masc. fem.*
- | | | | |
|------|-----------|-------|----------|
| bu | ‘father’ | omm | ‘mother’ |
| xu | ‘brother’ | oxt | ‘sister’ |
| ðkar | ‘male’ | antha | ‘female’ |

All inanimate nouns are assigned an arbitrary gender (3).

- (3) *Inanimates: masc. fem.*
- | | | | |
|--------|----------|-------|---------|
| babur | ‘boat’ | tawla | ‘table’ |
| korsi | ‘chair’ | nabta | ‘plant’ |
| ħebbak | ‘window’ | ħin | ‘eye’ |

¹In Arabic, the feminine ending is orthographically marked with a *ta marbuta* ‘tied *t*’. This ending *-t* is silent in spoken dialects except in *idhafa* constructions (construct state). For example, in *sayyara(t) Mariam* ‘Myriam’s car’, the *t* is pronounced. While some scholars and grammarians represent the *ta marbuta* as *-ah*, here the transliteration *-a* is used to better reflect its pronunciation in TA.

In many cases (mainly for suppletive and inanimates), gender is not morphologically marked on the noun. Nouns ending with the sound *-a* are often associated with the feminine gender, but this ending does not always signal that the noun is of the feminine gender. Many nouns ending in *-a* are masculine, and many feminine nouns do not end in *-a* (4).

(4)	<i>Masculine -a ending</i>	<i>feminine no -a ending</i>
	sma 'sky'	ʃams 'sun'
	dwa 'drug'	bit 'room'
	ʔomda 'mayor'	dar 'house'

The gender of nominals can be identified by agreement on controlled elements. Verbs and adjectives strictly agree in gender with the singular subject. This applies for both natural and arbitrary genders (5)-(6). The definite determiner *el* is invariant.

(5)	a.	El xal tkallam.	(6)	a.	emtiḥaan twiil
		the uncle.SG spoke.3MASC.SG			test long.MASC.SG
		'The uncle spoke.'			'a long test'
	b.	El xala tkallm-et.		b.	tawla kbir-a
		the aunt.SG spoke-3FEM.SG			table big-FEM.SG
		'The aunt spoke.'			'a big table'

In Modern Standard Arabic, plural agreement patterns differentiate gender. Consider the examples in (7), where both the adjectives and verbs inflect for gender in the plural.

- (7) a. Al-awlad al-jamil-uun katab-u. [Standard Arabic]
 DEF-boy.PL DEF-pretty-MASC.PL wrote-3MASC.PL
 ‘The pretty boys wrote.’
- b. Al-fatiy-at al-jamil-aat katab-na.
 DEF-girl-FEM.PL DEF-pretty-FEM.PL wrote-3FEM.PL
 ‘The pretty girls wrote.’

In the TA dialect, gender agreement is differentiated in the singular, yet with plural subjects, there is only one set of plural agreement forms, regardless of gender. To illustrate, *xwel* ‘uncles’ and *xalet* ‘aunts’ both trigger the same plural agreement marker *-u* on the verb in (8). In the same way, the plural noun *kressi* ‘chairs’ (masculine in the singular) and the feminine plural *twewel* ‘tables’ (feminine in the singular) are also followed by the unique plural adjective form *kbaar* ‘big’.

- | | |
|--|---|
| <p>(8) a. El xwel ketb-u.
 the uncle.PL wrote.3PL
 ‘The uncles wrote.’</p> <p>b. El xal-et ketb-u.
 the aunt-PL wrote-3PL
 ‘The mother wrote.’</p> | <p>(9) a. kressi kbaar
 chair.PL big.PL
 ‘big chairs’</p> <p>b. twewel kbaar
 table.PL big.PL
 ‘big tables’</p> |
|--|---|

These agreement patterns contrast with those found in Standard Arabic. Consider (10), which compares the suffixal agreement markers of Tunisian Arabic with those of Standard Arabic.

	STANDARD	TUNISIAN	
	1SG	- <i>tu</i>	- <i>t</i>
	2SG.MASC	- <i>ta</i>	- <i>t</i>
	2SG.FEM	- <i>ti</i>	- <i>t</i>
	3SG.MASC	-∅	-∅
	3SG.FEM	- <i>at</i>	- <i>et</i>
	1PL	- <i>na</i>	- <i>na</i>
	2PL.MASC	- <i>tum</i>	- <i>tu</i>
	2PL.FEM	- <i>tunna</i>	- <i>tu</i>
(10)	3PL.MASC	- <i>u</i>	- <i>u</i>
	3PL.FEM	- <i>na</i>	- <i>u</i>

Based on the patterns presented in (10), we observe that while Standard Arabic verbal agreement markers distinguish between all persons, genders and numbers, TA only distinguishes between genders in the third person singular, with ∅ in the masculine and *-et* in the feminine. Gender distinctions are neutralized in the plural in TA, and the suffix markers seem to converge towards the masculine (e.g., *-u* for the 3rd person plural), which is the default gender in Arabic. This convergence towards one specific form independent of gender is also found in plural suffixes on inanimate nouns.

There are two types of plural shapes in Arabic: the broken plural, which involves a change in the stem, and the sound plural, which is suffixational. [Chapter 4](#) discusses the different types of plurals in more details, and the way they are assigned to each noun. Here, I focus on the sound plural, since it is the only one that morphologically marks gender. The masculine plural suffix is *-een* (11-a)-(12-a), and the feminine plural marker is *-at*² (11-b)-(12-b).

²Sometimes *-et*, depending on the dialect. Here, *-at* and *-et* are used interchangeably.

Table 2.1: Gender assignment on Tunisian Arabic sound plurals

	ANIMATE	INANIMATE
MASCULINE	<i>-een</i>	<i>-at</i>
FEMININE	<i>-at</i>	<i>-at</i>

This, however, does not indicate whether plural inanimates are truly feminine or not. They trigger default agreement on verbs and adjectives (with no gender marking), like all other plural nouns in TA. In Standard Arabic, all non-human and inanimate nouns trigger feminine singular agreement on verbs and adjectives in the plural, regardless of their gender in the singular, and regardless of the form of the plural, as illustrated in (15).

- (15) Ja:ʔa-ti l-kila:bu. [Standard Arabic]
 came-FEM.SG DEF-dog.PL
 ‘The dogs came.’

(Fassi Fehri, 1988, 119)

This phenomenon is known as “deflected agreement” and is not observed in TA, as previously shown in (9). Deflected agreement is discussed in [Chapter 4](#). Plural inanimates trigger unmarked default agreement in TA and feminine singular agreement in Standard Arabic. Although in plural inanimates we cannot clearly distinguish the gender from agreement patterns, we can safely assume that they have some kind of feminine component. First, the fact that they take the plural suffix *-at*, that is also used to pluralize nouns denoting female entities, strongly suggests that plural inanimates are feminine. Second, feminine inanimates do infallibly trigger feminine agreement on verbs in Standard Arabic. This is a strong argument in favour of ana-

lyzing plural inanimates as feminine nouns. Therefore, in this work the plural suffix *-at* is considered as signaling the presence of the feminine on inanimates in TA, and not as a syncretic form.

To summarize the facts, the TA gender assignment system is best described as a “predominantly semantic” one (Corbett 1991, 13), where gender is based on sex for sex-differentiable nouns, and is arbitrary for the residue (Corbett, 2011). Both types of gender (natural and arbitrary) are reflected in agreement marking on verbs and adjectives. In the plural, only natural gender is preserved, while arbitrary gender on inanimate nouns is feminine by default. This indicates that gender in the plural depends on a diversity of factors, namely the gender assigned to the singular form, but also whether it is semantic or arbitrary. Based on these observations, it appears that gender is in some way connected to number in TA. Also, the notion of arbitrary gender on inanimates is blurry as assignment is completely arbitrary in the singular, but strictly feminine in the plural. It appears that gender is assigned based on sex of the referent for animate nouns, and partly based on number for inanimate nouns.

2.3 What is gender?

To tease apart the different manifestations and interpretations of gender, I first discuss the main definitions of gender available in the literature and how gender differs from other noun classification devices and ϕ -features. Then, I address the question of the locus of gender within the nominal spine. I argue that gender’s main function is that of nominalization, and hence, that it is hosted on the *n* head

(Lecarme, 2002; Kihm, 2005; Lowenstamm, 2008; Acquaviva, 2009; Kramer, 2009, 2014, 2015; Hammerly, 2018). This view is supported by the observation that a root, especially in a root and pattern language like Arabic, needs to combine with gender in order to yield a noun.

2.3.1 Defining gender

To account for the different functions and expressions of gender in TA, and eventually with other related languages, I shall first establish a definition for gender. The definition in (16) by Hockett (1958) has been adopted as the standard since the publishing of Corbett's (1991) influential book on gender.

(16) **Definition of gender (standard)**

Genders are classes of nouns reflected in the behavior of associated words.
(Hockett, 1958, 231)

Two elements stand out in this definition. First, gender is a way of classifying nouns. Second, gender involves agreement. These points provide avenues for exploration when it comes to studying gender. Some of the questions addressed in this section are: What are the different bases of classification of nouns (i.e., when not strictly semantic, are they always arbitrary)? What is the grammatical function of gender, when semantically vacuous? Are there any instances of gender that do not involve agreement? How does gender differ from noun classification?

In terms of gender assignment criteria, Corbett (1991) notes (crediting Hockett 1958, 231) that gender classification frequently corresponds to a real-word distinction

of sex, but does not have to, as “gender” etymologically derives from *genus*, via Old French *gendre*, and originally meant ‘kind’ or ‘sort’. So when not based on sex, what other principles govern the assignment of gender? Corbett (1991, 2011) and Comrie (1999) distinguish between two major factors: the semantic principle and the formal principle. According to the semantic principle, nouns are assigned a gender according to their meaning. According to the formal principle, nouns are assigned to gender according to their form. Gender assignment rules form a continuum from strict semantic systems Corbett (1991, 2011) to formal ones. Both ends of the continuum are very uncommon. On one hand, in many languages with semantic assignment systems, rules based on meaning do not completely cover the noun inventory. Corbett (1991, 2011) refers to these as “predominantly semantic assignment systems”. On the other hand, Corbett (1991) shows that there are no languages where formal assignment rules are sufficient to account for how gender is allotted to nouns. To illustrate, let us look at the case of Tamil and other Dravidian languages, considered to be strict semantic systems (Corbett, 2011, p.9). In Tamil, for example, nouns are divided into rational and non-rational. Rational nouns are divided into masculine and feminine, based on biological sex, while non-rational nouns are assigned the neutral gender. Such systems, also called “natural gender systems”, are very consistent. The speaker relies on the meaning of the noun in order to know its gender. Very few exceptions exist, where the gender must be learned. For example, the noun for ‘child’ in Tamil is *makavu* and is usually neuter (Corbett, 2011, p.9), but may also be masculine or feminine. It is interesting to look at how gender is assigned to nouns that fall outside the semantic rules. Corbett (1991, 2011) refers to such nouns as

“the remainder” or “the semantic residue”. In predominantly semantic systems, the remainder can converge towards one gender (e.g., neuter, masculine, feminine), or are shared between more than one gender. Amharic is an example of a system where the residue are assigned to one gender. Almost all inanimate nouns are masculine in Amharic (17), and only a handful are feminine (18) (Cohen 1970, 74, Leslau 1995, 161, Kramer 2015, 17).

(17) *Amharic: Inanimate masculine nouns* (from Kramer 2015, 18)

- | | | | | | |
|----|--------|----------|----|----------|---------------|
| a. | mot | ‘death’ | f. | wididdir | ‘competition’ |
| b. | kibir | ‘honor’ | g. | bet | ‘house’ |
| c. | wänbär | ‘chair’ | h. | dimmir | ‘total, sum’ |
| d. | dīngay | ‘stone’ | i. | wäräda | ‘district’ |
| e. | kibäb | ‘circle’ | j. | gazel’a | ‘newspaper’ |

(18) *Amharic: Inanimate feminine nouns* (from Kramer 2015, 18)

- | | | | | | |
|----|--------|-------------|----|----------------|----------|
| a. | mäkina | ‘car’ | e. | s’ähay | ‘sun’ |
| b. | azurit | ‘whirlpool’ | f. | kätäma | ‘city’ |
| c. | agri | ‘country’ | g. | betä kristiyan | ‘church’ |
| d. | midir | ‘earth’ | h. | tj’äräk’a | ‘moon’ |

When the remainder are distributed over more than one gender, it can be done based on formal properties (phonological or morphological), or in a completely arbitrary fashion. Corbett (2011) gives the example of Qafar, a Cushitic language

spoken in north-eastern Ethiopia and Djibouti, which has semantic assignment rules for animate nouns, and phonological assignment rules for the residue. Nouns that end with an accented vowel are feminine (for example, *karmà* ‘autumn’), while all others are masculine. In general, animate nouns follow these phonological rules (i.e., nouns denoting female entities tend to end with an accented vowel). However, when there is conflict, the semantic rules take precedence. Formal assignment rules can also be based on morphological information. This is the case in Russian (Corbett 1991, 33-34, Corbett 2011). For animate nouns, gender is assigned based on biological sex (male or female), while the residue are distributed over three genders: masculine, feminine and neuter. Russian has four main inflectional classes (Table 2.2). The assignment rules are straightforward: nouns in class I are masculine, nouns in classes II and III are feminine, and nouns in class IV are neuter.

Table 2.2: Inflectional classes in Russian (from Corbett 2011)

	I	II	III	IV
Nominative	<i>žurnal</i>	<i>gazeta</i>	<i>kost'</i>	<i>pis'mo</i>
Accusative	<i>žurnal</i>	<i>gazetu</i>	<i>kost'</i>	<i>pis'mo</i>
Genitive	<i>žurnala</i>	<i>gazety</i>	<i>kosti</i>	<i>pis'ma</i>
Dative	<i>žurnalu</i>	<i>gazete</i>	<i>kosti</i>	<i>pis'mu</i>
Instrumental	<i>žurnalom</i>	<i>gazetoj</i>	<i>kost'ju</i>	<i>pis'mom</i>
Locative	<i>žurnale</i>	<i>gazete</i>	<i>kosti</i>	<i>pis'me</i>
gloss	‘magazine’	‘newspaper’	‘bone’	‘letter’

Many of the sex-differentiable nouns belong to the class corresponding to their gender. For example, *mal'čik* ‘boy’ is in class I, and *sestra* ‘sister’ is in class II. It may be tempting to think we could dispense with semantic assignment rules, however, once

again, in conflicting cases, the semantic rules take precedence. For example, *djadja* ‘uncle’ denotes a male entity but is in class II, which is typically composed of feminine nouns. *Djadja* ‘uncle’ is masculine. This supports the idea that we do not find languages where formal assignment rules are sufficient. [Corbett \(1991, 2011\)](#) shows that gender assignment systems form a continuum between strict semantic to strict formal assignment, where both ends of the continuum are very uncommon. Even languages that are considered to have strict semantic gender assignment systems, such as Tamil and other Dravidian languages, have exceptions. While languages vary in whether the relevant semantic property is biological sex, animacy, humanness, or others, no language assigns genders to nouns completely randomly or completely formally. There is always a core set of nouns for which semantics determines how they receive gender. [Kramer \(2015, 70\)](#) notes that one of the most robust discoveries from typological research is that gender systems are always rooted in nominal semantics, and that all gender system have a “semantic core” (see also [Aksenov 1984](#), [Corbett 1991, 63](#), [Nichols 1992, 129](#), [Aikhenvald 2000, 25](#), [Dahl et al. 2000](#), [Corbett 2006, 752](#), [Matasović 2004, 22](#)). She therefore proposes the definition in (19) to reflect the key finding that all gender systems have a semantic core.

(19) **Kramer’s definition of gender:**

Gender is the sorting of nouns into two or more classes, depending on biological sex, animacy, and/or human-ness, for at least some animate nouns, reflected by agreement on other elements (e.g., adjectives, determiners, verbs, auxiliaries).

([Kramer 2015, 70](#))

The second important observation by [Hockett \(1958, 231\)](#) is that agreement plays a crucial role in defining gender (see also [Greenberg 1978, 49](#), [Nichols 1992, 124](#), [Comrie 1999](#), [Velupillai 2012, 165](#), [Kramer 2015, 65](#)). When compared to other noun classification devices, gender is distinguished by the fact that it is realized through agreement ([Dixon 1986](#), [Aikhenvald 2000, 19](#), [Grinevald 2000](#)), as opposed to noun classifiers, which are always associated with the noun itself ([Dixon 1986](#), [Aikhenvald 2000, 81](#)). In this work, I discuss the different types of noun classification systems. Different authors have different ways of categorizing such devices. [Aikhenvald \(2000\)](#) presents a fine-grained typology of noun categorization devices, outlining five broad types of systems, of which only three are relevant for this discussion: gender/noun class systems, noun classifiers, and numeral classifiers.

The properties of gender are the same as those of noun class systems. In fact, “nouns class” and “gender” are often used interchangeably, depending on the linguistic tradition. [Dixon \(1982, 1986\)](#) describes gender as a small finite set of classes, forming a closed grammatical system, where marking is never entirely within the noun word. [Aikhenvald \(2000, 19\)](#) echoes this sentiment in her definition of noun classes, where some constituent outside the noun itself must agree in gender with the noun. In some languages, in addition to the realization of gender through agreement, gender is also marked on the noun itself, while in other languages, nouns bear no overt markers. To sum up, while morphological marking on the noun is optional, the presence of agreement is the hallmark of gender.

Noun classifiers, however, typically co-occur with the noun in the noun phrase and do not participate in agreement processes. The choice of a classifier is based on

semantics and may correlate with inherent characteristics of nouns, such as the form of the referent or its type (animal, human, plant, etc.). All nouns in a language do not necessarily take a classifier, but may take one or several ones, unlike gender that is obligatory and unique for each noun (Aikhenvald, 2000, 81).

Finally, numeral classifiers have the same characteristics as noun classifiers, the only difference is that the former occur contiguous with numerals or expressions of quantity and are not required to appear on any other constituents (Aikhenvald, 2000, 98). In fact, Dixon's (1986) typology of noun categorization devices does not distinguish between nominal and numeral classifiers. He simply opposes classifiers with noun classes, arguing that the former is a grammatical phenomenon, while the latter is a lexico-syntactic one. In sum, while noun classes involve agreement, classifiers do not. Classifiers instead characterize the noun, with their choice being determined by lexical selection, and not by matching any inflectional properties of nouns with other constituents of a noun phrase.

To sum up, gender is a grammatical property of nouns and differs from other noun classification devices mainly because it is expressed on agreeing categories rather than on the noun itself. Gender is also obligatory on all nouns in languages with a gender system. This suggests that gender has a grammatical property that goes beyond semantic characterization of a noun. One of the aims of this chapter is to define this function.

2.3.2 Gender as a nominalizer

The first part of this section was concerned with the definition of gender and its grammatical characteristics. I now discuss its function and argue, following many others (Lecarme 2002; Kihm 2005; Lowenstamm 2008; Acquaviva 2009; Kramer 2009, 2014, 2015; Hammerly 2018), that gender has a nominalizing function.

One of the basic assumptions in DM is that the lexicon consists of roots that lack category and phonological form, and that may be semantically underspecified to varying degrees (see Marantz 1997). Roots acquire a category by being inserted in particular syntactic configurations, thus becoming morphemes. Morphemes are then associated with phonological features in the vocabulary component of morphology and may be called “exponents” or “Vocabulary Items” (Embick and Noyer, 1999). In what follows, I argue, following Lecarme (2002), Kihm (2005), Lowenstamm (2008), Acquaviva (2009), Kramer (2009, 2014, 2015), and Hammerly (2018), that gender is the exponent of *n*, the syntactic head responsible for turning roots into nouns in DM.

As defined in the previous section, gender is characteristic of nouns. Even when morphologically expressed on other categories, gender only reflects a feature that originated in nouns. We have already established that gender expresses some classificatory faculty of the mind, by having a semantic core (e.g., denoting biological sex in animate entities); however, gender assignment principles are never completely semantic. When the sex of an entity is not relevant, classification is done in an unsystematic way and gender seems to play no other role than defining word classes. This is the case for inanimate nouns in Romance, for example. For this reason, we

can conclude that gender is “useless” as a classificatory device. If classification was the sole function of gender, then its assignment would be based on clear principles. I have shown here through numerous examples, however, that this is not always the case. This has led different authors such as [Lecarme \(2002\)](#), [Kihm \(2005\)](#), [Lowenstamm \(2008\)](#), [Acquaviva \(2009\)](#), [Kramer \(2009, 2014, 2015\)](#), and [Hammerly \(2018\)](#), among others, to argue that the role of gender is to nominalize roots.

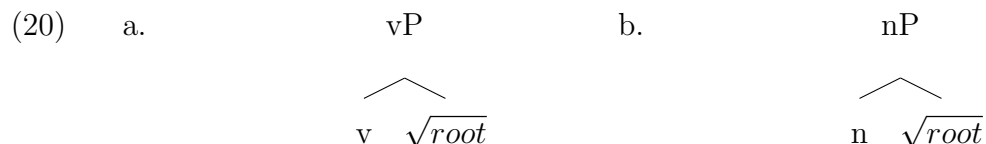
2.3.3 Gender in root and pattern languages ([Arad, 2005](#))

That gender is a grammatical device required for noun formation gets further support from root and pattern languages, where the root, as a set of consonants, has no fixed semantic or phonological realization prior to combination with word-creating morphology or patterns. As a root and pattern language, Arabic is conventionally analyzed as forming nouns from a consonantal root combined with one of a large number of nominal patterns. Let us take the notorious example of the consonantal root $/k t b/$, which refers to the basic concept of writing. This root can be combined with a verbal pattern, to give *kataba* ‘he wrote’, but also with one or more nominal patterns to form semantically-related nouns, e.g., *kaateb* ‘writer’ and *kitaab* ‘book’, as illustrated in Table 2.3.

Table 2.3: The Arabic root /k t b/ and its patterns

/k t b/	'to write'	
kitaab	'book'	MASC
kaatib	'writer'	MASC
maktab	'desk'	MASC
maktaba	'library'	FEM
maktuub	'written'	-
yaktab	'he writes'	-
kataba	'he wrote'	-

The examples in Table 2.3 show that vocalic pattern insertion not only turns consonantal roots into pronounceable strings, but also gives them a category and fixed semantic interpretation. It can be observed from these examples that when a root combines with a nominal pattern, gender is always spelled out. [Arad \(2005, ch.2\)](#) proposes that verbal patterns are inserted at the verbalizing head *v* so nominal patterns by analogy will be inserted at *n* (20).



Based on the observation of syllabic roots in Hebrew (mostly originated from borrowings that entered the language at an early stage), [Arad \(2005, ch.2\)](#) notes that only verbs require internal modification, while nominals do not. Syllabic roots contrast sharply with consonantal ones in their phonological requirement. They do not need to be altered by vocalic pattern insertion in order to become pronounceable.

This suggests that when vocalic patterns are inserted in a syllabic root, they reflect the spell-out of a syntactic rule, and not a phonological requirement. The list in (21) is composed of Hebrew nouns that are built from syllabic roots. These examples show that noun-creating morphology is non-overt with syllabic roots, as with typical English words (*cat*, *knife*, or *chair*).

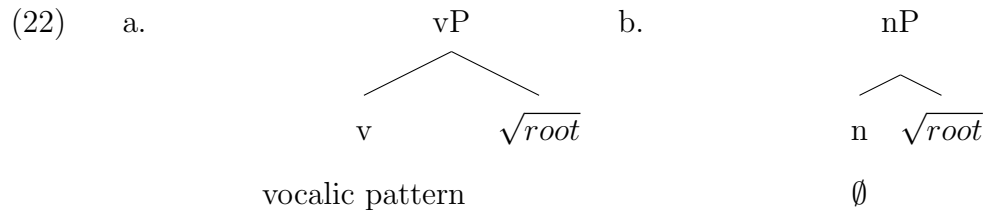
- (21) a. bdolax (crystal) d. zug (chair) [Hebrew]
 b. sus (horse) e. melafefon (cucumber)
 c. tarnegol (rooster) f. psefas (mosaic)
 (Arad 2005, ch.2)

This indicates that nominal patterns' only purpose is to make the consonantal root pronounceable. If they are not needed for phonological purposes, as is the case for syllabic roots, they will not be inserted. Assuming that verbal patterns have the same phonological function, it is expected that they will also not be necessary with syllabic roots. This, however, is not the case. Pattern morphology is obligatory for all verbs, including borrowed verbs with a syllabic root. For instance, this is the case of the following verbs in (22), made from nouns borrowed from English, Arabic, and Yiddish, respectively. The examples in Table 2.4 show that unlike nouns, verbs derived from syllabic roots must appear with vocalic patterns that are specific to Hebrew verbal morphology:

Table 2.4: Borrowed Hebrew nouns and verbs - [Arad \(2005, 35\)](#)

- a. telefon (telephone, n) tilfen (to telephone, v, pattern 3)
- b. mastul (drunk, adj.) hitmastel (to get drunk, v, pattern 7)
- c. xrop (a snooze, n) xarap (to snooze, v, pattern 1)

These data suggest that verbal patterns are more than a mere phonological requirement. Arad concludes from these observations that verbal patterns have a double purpose: they make the consonantal root pronounceable and provide the category feature. In sum, nominal patterns are inserted at PF in order to turn consonantal roots into licit phonological representations, unlike verb-creating morphology, which is the spell-out of a syntactic operation. In Arad's view, verb-creating morphology is reflected by vocalic pattern (22-a), while noun-creating morphology is non-overt (22-b).



For [Kramer \(2015\)](#), Arad's approach predicts that the choice of a nominal pattern in root and pattern morphology could be conditioned by gender. This prediction is borne out in Amharic (another root and pattern language), where some nominals vary in pattern according to gender (23).

	Root	Male: CiCuC	Female: CiCiCt
(23)	a. $\sqrt{\text{NGS}}$	nigus ‘king’	nigist ‘queen’
	b. $\sqrt{\text{L?L}}$	li?ul ‘prince’	li?ilt ‘princess’
	c. $\sqrt{\text{K’DS}}$	k’iddus ‘saint(m.)’	k’iddist ‘saint (f.)’

Each pair of nouns in (23) shares the same root, but has different patterns for natural gender. The male patterns have [u] as second vowel, whereas the female patterns have [i] as the second vowel and a *-t* suffix. Kramer assumes that the nominal pattern is inserted at *n* when overt³ and that it is capable of being conditioned by gender. This provides evidence that gender features are on *n*.

The data from TA corroborate the generalizations drawn from Hebrew (Arad 2005 and Kramer 2015). First, consider the examples from French borrowings in Table 2.5. We can see that nouns borrowed from French keep the same stem as the original French noun. However, verbs derived from those nouns obligatorily go through stem-internal modifications.

Table 2.5: Tunisian Arabic borrowings from French.

a. tilifun (telephone, n)	telfen (to telephone, v.)
b. duʃ (shower, n.)	idawweʃ (to shower, v.)
c. guigna (bad luck/boredom, n.)	tgayyen (to become boring, v.)
d. risk (risk, n.)	ireski (to take a risk, v.)
e. trafik (traffic, n.)	itarfek (to tamper with, v.)

I, therefore, draw the same conclusions for TA as Arad for Hebrew: nominal patterns are a phonological requirement, while verbal morphology must be spelled out to reflect features present in the syntax (i.e., voice, tense, aspect). TA also

³This treatment of nominal pattern is different from Arad’s, who considers nominal patterns to be inserted only at PF.

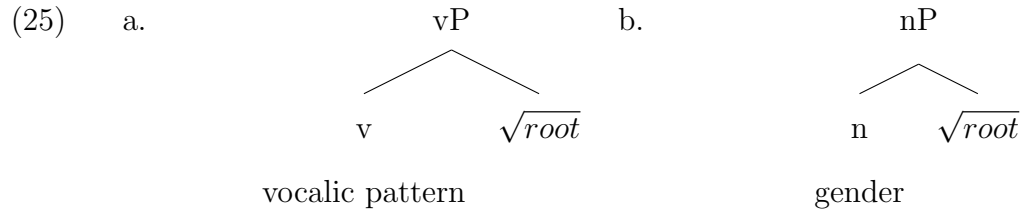
exhibits cases of nominal patterns that are conditioned by gender, as shown in (24), where each pair of noun shares the same root but has a different pattern for natural gender.

- (24) a. tfoll ~ tofla [Tunisian Arabic]
boy girl
'boy, girl'
- b. xu ~ oxt
brother sister
'brother, sister'

The conditioning of nominal pattern by gender shown in (24) provides additional evidence that gender features are on *n*, as previously argued by [Kramer \(2015\)](#) for Amharic.

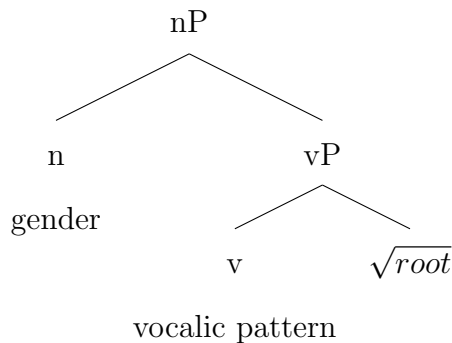
As another piece of evidence that gender is hosted on *n* based on root and pattern morphology, I add that even in the cases where the nominal pattern is non-overt (i.e., syllabic roots originating from borrowings), gender is always expressed on nouns, either on nominal morphology (via suffixation) or through agreement. Therefore, while distinctive nominal morphology can be optional - even on root and pattern languages like Arabic, which are notorious for exhibiting categorial information through vocalic patterns - gender is always present. This suggests that, if there were to be a distinctive mark of the nominal category, it should be the presence of gender rather than a specific vocalic pattern. Gender, unlike vocalic patterns, is not required by phonology, as evidenced by the examples in section 2.2 where gender is not overtly marked on nouns. This suggests that gender reflects the categorial identity of nouns, and,

therefore, represents syntactic operations. Whether or not gender is interpreted at both interface levels is discussed next. The temporary conclusion to draw from this is that gender is, more than vocalic pattern, the hallmark of the nominal category (25).



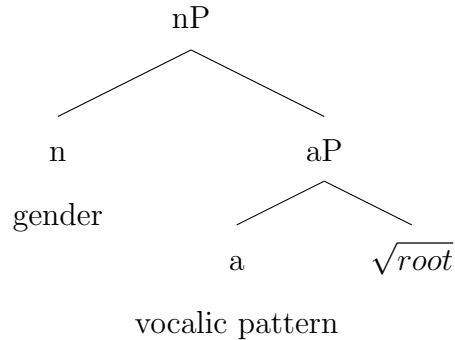
Another piece of evidence that gender is on *n* comes from nominalizations. In general, *n* is used to nominalize roots, as shown in (25-a), but also to nominalize other syntactic categories (Marantz, 2001; Arad, 2003, 2005). For example, a *n* can combine syntactically with a *vP* to create a deverbal noun (26), or with an adjective to form an abstract noun (27). Such derivations always involve gender (marked on the noun or through agreement).

- (26) a. /f r b/ feereb (to drink - drunk) [Tunisian Arabic]
 b. /f t ħ/ maftuħ (to open - open)



(27) saʕeed - saʕada (happy - happiness)

[Tunisian Arabic]



In this section, I have argued that gender features are even more closely linked to noun formation than nominal patterns in languages with root and pattern systems. These data provide more support to [Kihm's \(2005\)](#) claim that noun class morphemes in general (including gender) are exponents of *n* (also reflected in [Lecarme 2002](#); [Lowenstamm 2008](#); [Kramer 2009, 2014, 2015](#)), discussed next.

2.3.4 Class = *n* ([Kihm, 2005](#))

[Kihm \(2005\)](#) makes a distinction between gender and Class. He defines Class as a noun-forming device that is hosted on *n*, taking the root as a complement and hence giving it the suitable syntactic environment to turn it into a noun. Gender is a subtype of Class that is endowed with a classificatory (or semantic) content of varying richness or relevance depending on the language. For instance, when Romance word classes sort entities, they do so according to the criterion of biological sex, and this is what Kihm identifies as gender. He therefore assumes that in Romance, the classificatory feature (M/F) is distinct from the categorical feature N. The latter

feature is inherent to the word class morpheme, being its grammatical *raison d'être*: class morphemes form nouns from roots. The former feature, in contrast, is linked only when gender is relevant, given the meaning of the root to which the word class morpheme is attached. Thus, *gatta* ‘female cat’ and *porta* ‘door’ in Italian receive the following representations:

(28) $\sqrt{\text{gatt-a}}$ N(II) F

(29) $\sqrt{\text{port-a}}$ N(II)

Gatta is indeed feminine (classificatory feminine); *porta*, however, is only class II. The same is true for *gatto* when it denotes a male cat (N(I)M) and *libro* (N(I)). What remains is an entailment to the effect that Class I and II nouns whose roots denote relevantly sexed kinds refer to the masculine and feminine members of that kind respectively. Essentially, Romance animate nouns have both Class and gender, while inanimate nouns only have Class. Therefore, in Kihm’s terms, there is no such thing as arbitrary gender. This concept has been replaced by the term “Class” to reflect its nominalizing function and take away the false assumption that it somehow represents a classificatory faculty of the mind.⁴ In sum, Kihm argues that Class is the one possible content of the *n* head already proposed in the literature (see e.g., [Marantz 1997](#)) and that the differences observed in its position and degree of fusion

⁴Whether or not the *-a* in non-mating pairs belongs to the root does not matter here. In Kihm’s analysis, the ending vowel determines class in Romance languages, but we also know that Class/gender is not typically marked on the noun, as agreed upon in all definitions of gender discussed above. We could imagine an analysis where the ending vowel is part of the root, but the feature triggering agreement on controlled categories – call it Class or gender – would still be a feature on *n*, independently of how we interpret the ending vowel.

with the root are related to how contentful the head is. Put simply, Kihm assumes the presence of gender features only when Class is somehow semantically interpretable. Otherwise, we are simply dealing with Class and no gender features are involved. I adopt a slightly different terminology, where I suppress the notion of a supercategory called “Class” for numerous reasons. First, based on the definition of gender provided in [Section 2.3.1](#), I use this simple diagnostic to detect the presence of gender features: if a noun classification device triggers agreement on controlled categories (e.g., verbs, adjectives, etc.), it is gender. Whether or not this feature reflects a semantic property is addressed differently, in terms of feature interpretability. “Class” is a confusing term in this context, as it can also refer to other different related concepts like noun classes, classifiers, etc. [Aronoff \(1994, 64\)](#) for instance, defines Class as a set of lexemes whose members each select the same set of inflectional morphemes, while referring to gender as nominal agreement class. Therefore, I refer to Kihm’s larger notion of Class as *gender*, regardless of its level of semantic contribution, based on the simple observation that it is expressed through agreement. To account for the distinction between semantic and arbitrary gender, I adopt Kramer’s approach ([2009; 2014; 2015](#)), where gender features reside on *n* and can be interpretable or uninterpretable.

2.3.5 Different flavours of *n* ([Kramer, 2009, 2014, 2015](#))

Kramer’s proposal is based on data from Amharic, where the gender system is heavily reliant on sex distinctions for animate nouns, and most inanimate nouns and entities for which natural gender is unknown/irrelevant are masculine. There are,

however, a few feminine inanimates and feminine default animals. She accounts for this system with the following list of types on *n*:

- (30) **Types of *n* in Amharic (Kramer, 2015, 41)**
- a. *n* *i* [+FEM] Female natural gender
 - b. *n* *i* [-FEM] Male natural gender
 - c. *n* No natural gender (or natural gender irrelevant/unknown)
 - d. *n* *u* [+FEM] Feminine arbitrary gender

The interpretable types of *n* in (30-a) and (30-b) are for animate entities with natural feminine or masculine gender. Nouns referring to inanimate entities or animate entities with unknown or irrelevant gender are masculine by default, as represented in (30-c). The few exceptions where inanimate nominals are feminine are licensed under an uninterpretable [+FEM] *n*, as in (30-d).

Kramer's proposal is based on the idea that feature valuation and interpretability (two dimensions of agreement) are independent concepts (Pesetsky and Torrego, 2004). Pesetsky and Torrego (2004) abandon Chomsky's Valuation/Interpretability Biconditional which stipulates that a feature *F* is uninterpretable iff *F* is unvalued (in this sense, valuation is a lexical encoding of interpretability). The elimination of the Biconditional allows lexical items to come from the lexicon with features that display two combinations of properties not accepted by the MI/DbP framework: (i) uninterpretable but valued; and (ii) interpretable but unvalued.

Recall from Section 2.2 that masculine is the default gender in TA, as reflected by agreement triggered by DPs where gender is unknown/irrelevant (31). This is

why I have identified the feminine as being the marked feature.

- (31) a. ʃkʊn xlat l-owwel? [Tunisian Arabic]
who arrived.MASC.SG DEF-first.MASC.SG
'Who arrived first?'
- b. had ma xrej.
person NEG left.MASC.SG
'No one left.'

Unlike in Amharic, Arabic inanimate nouns are not relegated to the default masculine; they are equally distributed among the feminine and the masculine genders (32). Nouns referring to animate entities of an unknown or irrelevant gender tend to be masculine, but there are some exceptions (33).

(32) **Feminine inanimates in TA**

- a. ʃams (sun)
- b. sma (sky)
- c. tawla (table)
- d. marqa (stew)

(33) **Feminine animate of unknown/irrelevant gender in TA**

- a. naħla (bee)
- b. rethla (spider)
- c. zarrafa (giraffe)
- d. om el buya (owl)

Based on the arguments provided above, by supporting the idea of gender being an exponent of the n head, and also to account for natural/arbitrary gender distinction in Arabic, I adopt Kramer’s system with the features in (34).

(34) **Types of n in TA**

- a. n i [+FEM] Female natural gender
- b. n i [-FEM] Male natural gender
- c. n No natural gender (or natural gender irrelevant/unknown)
- d. n u [+FEM] Feminine arbitrary gender
- e. n u [-FEM] Masculine arbitrary gender

As for Amharic, interpretable [+/-FEM] n is reserved for animate nominals where gender reflects biological sex. n with no gender feature is reserved for animates with unknown/irrelevant gender; these nouns are masculine by default. Feminine animate nominals with an unknown/irrelevant natural gender are licensed under an uninterpretable [+FEM] n . Finally, inanimate nouns are assigned arbitrary genders and are licensed under uninterpretable [+/-FEM] n .

To round out the discussion, I now address gender licensing in different-root nominal pairs such as *bu* ‘father’ and *omm* ‘mother’. It seems that each of these roots have a “male” or “female” component as part of their meaning. Therefore, it may be tempting to include gender as a part of the information encoded in the root. There is, however, a way of accounting for different root nominals while keeping the roots free from features that are associated with particular categories, such as gender. This approach also has the advantage of keeping all gender features on

the same head, namely n . Licensing conditions are mechanisms that match root and category-defining heads in Distributed Morphology (see [Harley and Noyer 1998, 1999, 2000](#); [Galani 2004](#); [Siddiqi 2009](#)). [Acquaviva \(2009\)](#) proposes that gender is assigned on the basis of licensing conditions that determine which roots combine with which gender. In this perspective, saying that “a noun has gender X” means “a root Vocabulary item is licensed in the context of [n] with gender X”. [Kramer \(2009, 2014, 2015\)](#) takes this account further and adapts it to her model by stipulating that each Amharic different-root nominal is licensed under one of either $n[+FEM]$ or $n[-FEM]$ (35).

(35) **Semantic licensing condition: ‘mother’** (adapted from [Kramer 2015](#), 51)

$[n[+FEM] [\sqrt{\text{omm}}]] = \text{‘female parent’}$

Because these licensing conditions affect interpretation, Kramer proposes that they are encoded in the Encyclopedia as conditions on the semantic interpretation of a root in a context. For example, $\sqrt{\text{omm}}$ ‘mother’ is only interpretable at the Encyclopedia under a $n[+FEM]$. If the Encyclopedia receives any other n combined with $\sqrt{\text{omm}}$, it will be unable to interpret the structure and the derivation will crash. Same root nominals, on the other hand, are licensed under any n , as illustrated in (36), where the root combined with an n of the type $i[+FEM]$ makes a female artist (36-a), and the same root combined with an n of type $i[-FEM]$ makes a male artist (36-b).

(36) **Semantic licensing condition: ‘artist’**

- a. $[n[+FEM] [\sqrt{\text{fannen}}]] = \text{‘female artist’}$
- b. $[n[-FEM] [\sqrt{\text{fannen}}]] = \text{‘male artist’}$

In the analysis outlined here, there is no need to separate a gendered piece of meaning from the root. Licensing a root in a particular nominal context is what gives it a male or female interpretation; there is no inherent male-ness or female-ness to the roots themselves. It is true that natural gender (femaleness/maleness) correlates with being licensed under an interpretable $n[+FEM]$ or $n[-FEM]$, but that is because these features themselves are what trigger female/male interpretation, presumably at LF before the Encyclopedia is accessed. Morphologically, these features then trigger feminine exponents (for $n[+FEM]$) or masculine/default exponents (for $n[-FEM]$). The crucial difference with this perspective on gender licensing is that meaning arises in a construction, not a root.

The last category of nominals to be discussed is inanimate nouns. We have established that their gender features are irrelevant for semantic interpretation. However, in addition to ensuring that their roots are licensed under the right types of n , we must also make sure that other ns (e.g., interpretable ns) cannot combine with them. In line with root licensing conditions as discussed above, I adopt [Kramer’s](#) (2009, 2014, 2015) solution which consists of assuming a semantic licensing condition, ensuring that the root for inanimate nouns such as ‘table’, ‘bed’, ‘idea’, etc. can only be interpreted in the context of a n that lacks interpretable gender features (cf. [Harley and Noyer 1998, 1999, 2000](#)), as in the root of a noun carrying a “feminine” piece

of meaning can only be licensed under an interpretable [+FEM] *n*. This is shown for the TA root *dar* ‘house’ in (37).

- (37) **Semantic licensing conditions: ‘house’** (adapted from [Kramer 2015](#), 53)
[*u n*[+FEM] [$\sqrt{\text{dar}}$]] = ‘house’

Essentially, this licensing condition conveys the idea that a house is not the type of entity in the real world that can ever have natural gender.

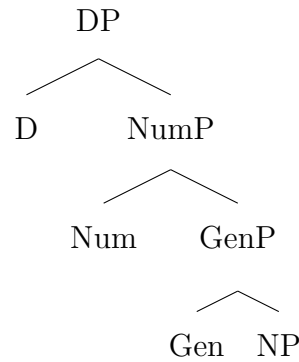
I now turn to addressing the default assignment of the feminine gender to all inanimate plurals in Arabic, in the light of the definitions of gender provided above and the analysis sketched out in the current section. I show that an *n*-based analysis of gender, although desirable for some of the data, does not account for the whole gender system in TA. I discuss the possibility for gender to be marked on another functional head.

2.4 Gender/number dependencies in TA

2.4.1 Certain assumptions about the structure of the DP

Based on data from Catalan, [Picallo \(1991\)](#) argues that gender heads its own functional projection, GenP, immediately dominating NP. GenP is the source of gender inflection for all nominals and is dominated by NumP, which contains number inflection, to give the structure represented in (38).

(38)



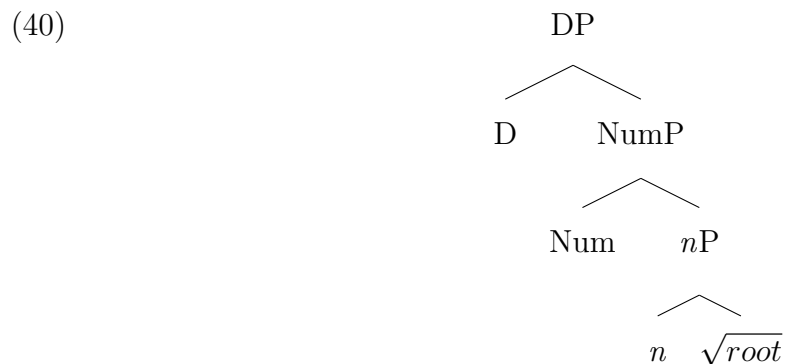
This structure successfully predicts gender inflection to be closer to the stem than Num, as evidenced by data from Romance languages such as Spanish, where the plural morpheme *-s* is added after the gender morpheme (*-o/a*) (39).

- (39) a. niñ-o-s (boy-m-pl) [Spanish]
b. niñ-a-s (daughter-f-pl)

However, following the arguments put forth in the previous sections, the same predictions can be achieved concerning this data without the presence of GenP. Gender morphology is also closer to the stem than number morphology if gender is simply put on *n*. Moreover, [Kramer \(2015\)](#) points out that GenP would be a projection that has no consistent semantics. Arbitrary gender on inanimate nouns, for instance, is uninterpretable, and [Chomsky \(1995, 349-55\)](#) argues against including in the syntax projections that only contain uninterpretable features (e.g., Agr nodes). Therefore, if gender on *n* accounts for the nominalizing function of gender, I argue against the inclusion of a gender functional projection in the syntax. Instead, I follow accounts where gender can be fully accounted for as a feature either on *n*, or, as I argue in

Section 2.5, on Num (Ritter, 1993).

Furthermore, I follow Ritter (1991, 1993) and many others (Farkas, 1990; Alexiadou, 2004) in assuming a functional projection Num bearing number specification of the noun phrase. In Chapter 3, I discuss another possible role for the Num head, namely, the function of division as put forth by Borer (2005)⁵ and many others following her work. As discussed in the previous section, I adopt the DM framework, where the root is underspecified in terms of category and must combine with a functional head in order to obtain a category, resulting in the structure in (40).



2.4.2 Gender is conditioned by number in TA inanimates

Returning to the data from TA, let us review the emergence of gender and number features in TA. First, gender does not need to be marked on singular nouns, and no overt morpheme marks the singular. Consider, for example (41-a), where gender is marked on the feminine members of mating pairs, but not on other nouns, whether animate (41-b) or inanimate (41-c). In all cases, no overt morpheme is associated to the singular.

⁵In Borer's work, the number projection is dubbed Div (division).

- (41) a. fannen - fannen-a ‘artist (m.), artist (f.)’ [Tunisian Arabic]
 b. xu - oxt ‘brother, sister’
 c. babur - nabta ‘boat (m.), plant (f.)’

Plural suffixes, however, always inflect for gender, as in (42).

- (42) fannen-*een* - fannen-*at* ‘artists (m.), artists (f.)’

For inanimate nouns, all plural suffixes take the feminine shape *-at*, regardless of their gender in the singular.

- (43) babur-*at* - nabt-*at* ‘boats, plants’

The main observation here is that only interpretable gender is preserved in the plural. On the basis of this data, there are two alternative ways to analyze plural suffixes in TA. The first option would be to consider *-een* and *-at* as portmanteau morphemes for gender and number, that is, exponents for the fused *n* and Num heads. The second option is to consider *-een* and *-at* as plural allomorphs that are conditioned by the content of *n*. In what follows, I explore both options and discuss their shortcomings, before pursuing a third one that is empirically superior.

Fusion is a post-syntactic operation where two sister terminal nodes are fused into a single terminal node before Vocabulary Insertion, resulting in an exponent containing the features of both nodes (Halle and Marantz 1993, 136, Halle and Marantz 2004, 277, Halle 1997, 148). The *n* head and the number head are adjacent, and thus are local enough to be merged into a single item. This would result in an operation

such as (44), where the left part represents discrete features, and the right, spelled out exponents at PF. In the case of TA, the portmanteau affixes would have the representation in (45).

- (44) [n:X] [Num:Y] → [n:X, Num:Y]
- (45) a. [n:i[-FEM]] [Num:PL] → [-een]
 b. [n:i[+FEM]] [Num:PL] → [-at]
 c. [n:u[+/-FEM]] [Num:PL] → [-at]

Similar analyses have been put forth for Italian plural affixes by [Alexiadou \(2004\)](#) and [Acquaviva \(2009\)](#). Unlike Spanish, Italian lacks a single plural marker. Instead, the distribution of number affixes is described as in (46).

		SINGULAR	PLURAL
(46)	MASCULINE	-o	-i
	FEMININE	-a	-e

[Alexiadou's \(2004\)](#) interpretation of the above paradigm is that the terminal vowel is a portmanteau morpheme for class/gender and number, since the choice of plural suffixes for Italian nouns is based on gender. Along the same lines, [Acquaviva \(2009\)](#) claims that gender and number always have a fused exponence in Italian, since there is no discrete plural morpheme comparable to the Spanish or English -s, which expresses number independently on the noun's gender. If gender and number features are hosted on different heads, *n* and the immediately higher Num, the two heads must undergo Fusion as a property of Italian morphology. Applying this analysis to the

TA data, the suffixes *-een* and *-at* would be viewed as portmanteau morphemes for gender and number, and not just plural markers.

Bernstein (1993) argues for an anti-decomposition analysis of similar data in Walloon, where the feminine and masculine plural markers (*-ès* and *-s*, respectively) are analyzed as complete units and not complex morphemes. This view is unproblematic for the masculine *-s*, since we are dealing with a single morpheme. In the feminine, however, it is not obvious that the affix is not morphologically complex, that is, *è* could be viewed as representing gender and *-s*, marking number. Nevertheless, *-è* is never found in singular feminine forms, and Bernstein argues on the basis of this, that plural markers are undecomposable elements in Walloon. Ferrari (2005) argues for a similar analysis of Italian plural markers, where the affixes *-e* and *-i* are considered as number markers, and *-a* and *-o* are viewed as gender markers only spelled out in the singular. She assumes that in Italian, as it is the case in Spanish, there is only one plurality marker, namely *-i*. The feminine plural marker *-e* can be conceived as the result of a phonological fusion between *-a* and *-i*, both syntactically present at the moment of plural formation. This leads us to the discussion of another alternative way of analyzing TA plural affixes, namely, contextual allomorphy.

According to the Fusion analysis, *-een* and *-at* in TA are the fused exponents of two distinct morphological features, gender and plural. *-een* is the exponent of the masculine and plural features, and *-at* is the exponent of the feminine and plural features. Another way to approach this distribution is to view these suffixes as plural allomorphs conditioned by the gender of *n*. In other words, *-een* and *-at* would be analyzed as number exponents rather than portmanteau for gender and number.

Num and *n* are adjacent in the nominal structure. Therefore, they are likely to condition each other allomorphically, that is, by Num being realized as a particular exponent in the context of a particular *n*. Conditioned allomorphy involves a choice among alternative Vocabulary Items that differ in their substantive morphosyntactic and phonological features (Halle and Marantz 1993). In the case of TA, the content of *n* (gender and interpretability) would serve as the context triggering allomorphy of the number morpheme, as illustrated in (47).

- (47) [PL] ↔ *-een* / [*n* *i*[-FEM]]
 [PL] ↔ *-at* / [*n* *i*[+FEM]]
 [PL] ↔ *-at* / [*n* *u*[+/-FEM]]

Moreover, assuming that spell-out is cyclic, the more deeply embedded morphemes are spelled out first. It would therefore make sense to consider that information about gender, which is closer to the root, be accessed for the spell-out of number, which is more outward.

Both Fusion and contextual allomorphy are plausible ways of describing the plural affixes in TA. When it comes to glossing TA sound plural nouns, one may include or exclude gender marking from the plural affix. A portmanteau/fusion analysis for TA plural suffixes does not, however, explain the change in gender in the plural for inanimate nouns. Given that an inanimate noun acquires a masculine or feminine gender for nominalization purposes, fusion does not explain where the feminine feature originates from in the case at hand. We can assume that *-at* is the Vocabulary Item for an uninterpretable *n* and a Num head bearing the plural feature, as in (48).

$$(48) \quad [n:u[+/-FEM]] [\text{Num:PL}] \rightarrow [-at]$$

This fusion operation does not faithfully translate the gender and number bundle into the expected morpheme in cases where the noun is masculine in the singular. We can assume that the unexpected plural form is the result of an uninterpretable gender on n and a plural feature on Num. Yet, based on the exponent coinciding with that of the plural of female referring nouns and the fact that feminine agreement is triggered in Standard Arabic, a feminine feature must be present somewhere.

The same issue arises if we consider the contextual allomorphy analysis. In this case, $-at$ would be considered as bearing only number features, but would have to be conditioned by the presence of the feminine gender on n . We can assume that the plural suffix $-at$ is conditioned by the presence of an uninterpretable gender on n (be it masculine or feminine), as illustrated in (49).

$$(49) \quad [\text{PL}] \leftrightarrow -at / [n \ u[+/-FEM]]$$

However, this view still assumes that the $-at$ exponent pluralizing inanimates is not feminine, yet results from the plural suffix being conditioned by the uninterpretable gender on n . None of these analyses account for gender switch in the plural of inanimates.

To keep all gender features on n , let us consider an option where the gender of inanimate nouns depends on the feature on Num; if the feature on Num is [PL], then this would condition uninterpretable n to take the feminine gender. This analysis, however, is not viable, mainly for two reasons. First, if gender on n depends on the

feature on Num, then why would it be variable in the singular? Second, assuming that spell-out is cyclical, the idea that exponence proceeds from the root outwards means that each spell-out operation has access to the more embedded nodes. Since n is closer to the root than Num, then this analysis can be ruled out.

An important generalization about how gender is projected on the Num phrase according to its interpretability comes from the observation that only interpretable gender is preserved at the higher Number level. There is a link to be made with animacy hierarchy: animate nouns normally get their gender from semantic rules, whereas inanimate nouns may or may not have semantic gender. It seems that the principles governing gender assignment on inanimates are in general weaker than the rules applying to animates (Dahl et al., 2000). There is a tendency for semantic agreement features to override grammatical ones, as noted by Corbett (1991), and as evidenced from mixed agreement patterns of hybrid nouns in Russian and German (see Steriopolo and Wiltschko 2010; Matushansky 2013; Landau 2015). For example, Steriopolo and Wiltschko argue that when natural gender is specified for a noun, the gender of the DP must be identical to it in order to avoid semantic conflict. Based on these claims that animate gender is “stronger” than inanimate gender, and is thus preserved across all numbers, let us consider an impoverishment analysis that would apply only to inanimate plurals. Impoverishment rules remove a feature from a syntactic feature bundle, resulting in the insertion of a less specified item instead (Bonet 1991; Noyer 1998; Harley 2008; Nevins 2011), as illustrated in (50).

(50) $[+/-\text{FEM}; +\text{PL}] \rightarrow [+ \text{PL}]$

While it is true that there is some kind of neutralization of gender distinctions in the plural for inanimate nouns, this is done to the benefit of the feminine marker. We know independently that masculine is the default gender in Arabic, with the feminine being more marked. Therefore, impoverishment is not an option.

To sum up, I concluded that *-at* truly signals the presence of a feminine feature in the derivation of inanimates. However, while I argued that gender is on *n* because it has a nominalizing function, I showed that feminine gender on *-at* cannot originate from *n*. In the next section, I show that gender can be hosted on Num in languages showcasing gender switch in the plural (Farkas 1990; Ritter 1993; Giurgea 2008; Croitor and Giurgea 2009). I argue that gender marking on inanimate plurals does not have a nominalizing function in TA. It is therefore hosted on Num due to its dependency on plural number. This analysis paves the way for my analysis of the singulative in [Chapter 3](#), where I show that gender on singulative nouns is interpretable on the Num head.

2.5 Proposal: uninterpretable gender on Num

In this section, I pursue the idea that gender can be, in some cases, expressed on Num (following Ritter 1993). For these cases, gender does not have a nominalizing function, but is rather an inherent feature of the Num head. This idea is also proposed by Farkas (1990) for Romanian and Bernstein (1993) for Walloon. I begin by summarizing four proposals pursuing the same idea using data from Romanian (Farkas 1990; Ritter 1993; Giurgea 2008) and Walloon (Bernstein 1993).

2.5.1 Previous Num-based analyses of gender

Farkas (1990) addresses the case of Romanian neuter nouns, which trigger masculine agreement in the singular, and feminine agreement in the plural, as in (51).

- (51) a. Un scaun confortabil e folositor. [Romanian]
a.MASC.SG chair comfortable.MASC.SG is useful.MASC.SG
'A comfortable chair is useful.'
- b. Niște scaune confortabile e folositor.
some chairs comfortable.FEM.PL are useful.FEM.PL
'Some comfortable chairs are useful.'

First, an analysis of the data in (51) must distinguish between nouns that are [-FEM] independently of number, and neuter nouns, which are [-FEM] only in the singular. Second, the analysis requires a rule that captures the observation that the gender of neuter nouns is predictable from their number. Farkas assumes a theory of underspecification in which the gender features of masculine nouns are lexically specified, and those of neuter nouns are introduced by a Feature Co-occurrence Rule (FCR). Such rules express connections between values of co-occurring features by predicting the value of a particular feature, given the value of some other feature. They generally resemble (52), where F and G are feature names and x and y are feature values:

$$(52) \quad Fx \rightarrow Gy$$

The FCR for neuter Romanian nouns triggering feminine agreement in the plural is presented in (53).

(53) **Feature Co-occurrence Rule for Romanian neuter nouns:**

$$[+PLURAL] \rightarrow [+FEM]$$

FCRs are strictly feature-filling rules, that is, they only operate on elements that lack a certain feature. Therefore, (53) applies only to neuter nouns, since they are not lexically specified for gender.

To capture the fact that singular nouns are masculine and assuming that [-FEM] is the default gender, a Feature Specification Default rule (also a feature-filling rule) is introduced. A default rule fills in the default value of a feature that has not been previously specified by some other principle. The default rule has the general form in (54), taking the specific form in (55) for Romanian neuter nouns.

$$(54) \quad [] \rightarrow Fx$$

(55) **Feature Specification Default for Romanian neuter nouns**

$$[] \rightarrow [-FEM]$$

The Elsewhere Principle ensures that (55) takes precedence over (53).

In sum, the FCR rule links up the [+PL] feature with the [+FEM] one, and the co-occurrence of the masculine gender with singular nouns is captured by the FSD rule, which accounts for the notion of unmarked, or “elsewhere” value. Farkas also notes that under the assumption that [-PL] is the default (unmarked) number value, the markedness of the gender feature of Romanian neuter nouns parallels that of their number feature. A similar parallel has been observed by [Pelletier \(1989\)](#) in three Dravidian languages (Telugu, Kurukh, Malto), where gender markedness correlates

with number markedness.

Bernstein (1993) looks at the case of the plural marker *-ès*, which only applies to feminine nouns. As previously discussed, she analyses *-ès* as a plural marker, and not as a complex morpheme that can be decomposed into a feminine affix (*-è*) and a plural one (*-s*), as *-è* alone never surfaces in feminine singular nouns. Moreover, data from Morin (1986) reveals that feminine marking on adjectives is only overt when the controlling noun is plural. The adjective *black* in (56) only contains a final *-e* when preceding a feminine plural noun (*currants* in (56-b)), but not when preceding a feminine singular (*thorn* in (56-a)).

- (56) a. li neûr sipène [Walloon]
the black thorn
'the black thorn (FEM.SG)'
- b. li nêure gruzales
the black currants
'the black currants'
- (Morin, 1986)

The data in (56) illustrate another case where gender features are contingent on number.

Ritter (1993) proposes a non-unified treatment of gender features, arguing that their syntactic location is subject to cross-linguistic variation. In particular, she claims that gender is a feature on N in Hebrew, but a feature on Num in Romance languages. Her hypothesis is supported by the analysis of irregular plurals in the studied languages. First, she shows that gender switching is a fairly productive

- c. yelad-ot nexmad-ot
 boy(FEM)-PL nice-FEM.PL
 ‘nice girls’

While the data in (58) and (57) suggest that number markers are gendered, it can be argued that the existence of two distinct plural forms show that gender is specified both on the noun stem (N) and on the plural suffix. However, this proposal makes two predictions that are not borne out in Hebrew. The first prediction is that feminine nouns will always take a feminine plural suffix and a masculine noun will always take a masculine plural suffix. The second assumption is that it will be the gender of the plural suffix, rather than the gender of the stem, that determines the gender of the derived form. Hebrew has masculine nouns that exceptionally select the feminine plural as well as feminine nouns that exceptionally select the masculine plural. In these cases, the gender of the stem, not the plural marker, is what triggers agreement on adjectives and verbs, as shown in (59).

- (59) a. šana tov-a [Hebrew]
 year(FEM) good-FEM
 ‘good year’
- b. šan-im tov-ot
 year(FEM)-PL good-FEM.PL
 ‘good years’

The noun *šana* ‘year’ is feminine in the singular, as evidenced by the agreement it triggers on the adjective (59-a). Although its plural form takes the so-called masculine suffix *-im*, it still triggers feminine agreement on the adjective and is

hence analyzed as feminine (59-b). This indicates that gender is on N, despite the fact that the plural marker corresponding to a different gender may be attached to a noun.

The situation is different with nouns in Romance languages. First, it appears that in these languages, gender switching is not a derivational strategy for deriving new nouns. Rather, it is a productive, inflectional operation used on animate nouns to express biological sex distinctions, as shown in (60) from Spanish ([Harris, 1991](#)), with pairs of nouns that have distinct but related masculine and feminine forms.

- (60) a. muchach-o muchach-a ‘boy, girl’ [Spanish]
b. jef-e jef-a ‘male chief, female chief’
c. poeta poetisa ‘poet, poetess’
([Harris, 1991](#))

Only a few pairs of inanimate nouns differ only in gender/word class marker in Spanish. Gender thus seems to operate like an inflectional feature in Spanish nouns, such as number. Ritter infers from these data that gender switching is not a feature on N in Romance languages. Further support for this claim comes from the data from Romanian ([Farkas, 1990](#)) and Walloon ([Bernstein, 1993](#)), as discussed above. From the Romanian data and based on Farkas’ analysis, Ritter concludes that it is the plural marker itself that is feminine in the case of neuter nouns. In other words, irregular plurals (of neuter nouns) in Romanian involve gender switching because Num, not the noun stem, bears gender specification of the noun phrase. Ritter also addresses the case of the suffix *-ès* in Walloon, and its analysis by [Bernstein](#)

(1993), concluding that *-ès* is analyzed as the overt realization of the functional head Num, as a number marker with inherent gender features. Based on the data from Spanish, which indicates that gender is an inflectional strategy, and from the behaviour of irregular plurals in Romanian and Walloon, Ritter concludes that in Romance languages, gender is found together with the number specification on a functional head, namely Num. This contrasts with Hebrew, where gender is on N, given its derivational function.

A Num-based analysis for Romanian neuters has also been proposed by Giurgea (2008) (also see Croitor and Giurgea 2009), in an attempt to formalize Corbett's (1991) distinction between *controller gender* or *nominal class* and *target gender*. This distinction was originally proposed for various systems, including Romanian, where the value of gender on targets of agreement depends on the value of the number feature. Assuming that the number feature is generated on a functional head Num, Giurgea (2008) proposed that in languages where gender depends on number, the gender feature is also generated on Num (as previously proposed by Ritter 1993 for Romance languages). The relation between gender and agreement class is mediated by selection, considering that it is generally assumed that functional categories such as Num select lexical categories. For example, suppose that Num comes in different flavours: [+SING, +MASC], [+SING, +FEM], [+PL, +MASC], [+PL, +FEM]. Each of these heads is specified for a given nominal class, which we may call Class I ("controller-masculine"), Class II ("controller-feminine") and Class III ("controller-neuter").

- (61)
- Num [SING, MASC] selects for class I and class III nouns
 - Num [SING, FEM] selects for class II nouns
 - Num [PL, MASC] selects for class I nouns
 - Num [PL, FEM] selects for class II and class III nouns

In all the systems of this type described by [Corbett \(1991\)](#), gender and number are fused under a single morpheme. This provides more evidence for gender and number being bundled on the same head.

2.5.2 A Num-based analysis of gender for TA inanimates

Returning to the data from TA, it appears that gender can be both analyzed as derivational and inflectional. On one hand, gender is responsible for forming nouns out of roots in Arabic, even more than vocalic patterns. On the other hand, the presence of mating pairs like *fannen-fannena* ‘male/female artist’ indicates that gender can also have an interpretive function in TA. To account for this distinction, I adopt [Kramer’s \(2009; 2015\)](#) analysis, where gender is hosted on *n*, by virtue of its nominalizing function, but with the specification *i/u*, depending on whether or not the gender is interpretable. This, however, does not suffice to explain why all plural inanimate nouns are feminine in TA. To account for this observation, and based on the four analyses discussed above, I argue that Num bears gender features in TA, and that they can override the gender features on *n*, depending on the nature of the latter.

First, an important observation is that interpretable features on *n* are always

preserved in the plural. This indicates that interpretable (semantic) gender cannot be overridden, regardless of the features present on Num. For the plural or inanimate nouns, I adopt an analysis similar to that of Farkas (1990), where the feminine plural marker is the result of a Feature Co-occurrence Rule (FCR). Remember that FCRs are feature-filling, which means that they only operate on elements that lack a certain feature. In terms of gender, the animacy hierarchy states that the propensity of gender being referential, rather than lexical, is increased as we go up to the animacy hierarchy (Dahl et al., 2000). Following this principle, I assume that interpretable gender on n is more likely to determine the gender of the whole DP. Based on this assumption, I consider that interpretable gender on n percolates to Num, and that gender feature on Num is hence specified for gender. This explains why animate nouns always take plural markers that match their gender. For inanimate nouns, gender on n is uninterpretable. I assume an impoverishment rule that removes the uninterpretable gender feature from the syntactic feature bundle in the context of a [+PL] Num head, as illustrated in (62).

$$(62) \quad [+/ \text{FEM}; +\text{PL}] \rightarrow [+PL]$$

Due to this impoverishment rule, gender on Num is unspecified for inanimate nouns. Following Farkas (1990), I assume an FCR that ensures the co-occurrence of the [+PL] feature with the [+FEM] feature, such as the case for Romanian (63).

(63) **Feature Co-occurrence Rule for TA inanimate nouns:**

$$[+\text{PLURAL}] \rightarrow [+FEM]$$

For singular nouns, the impoverishment rule defined in (62) does not apply, since it only applies in the context of [+PL]. Therefore, in the context of singular inanimates, it is the gender on *n* that percolates to Num. This analysis differs from that originally proposed by Farkas, since the gender of inanimates is only partially predictable by their number in TA, that is, only in the plural. The gender of Romanian neuter nouns, on the other hand, is always contingent on number, with all singulars being masculine and all the plurals being feminine. Therefore, I assume the gender of singulars to be lexically specified on *n* in TA, by stating that the impoverishment rule only applies to plural affixes. This proposal also aligns morphologically-marked gender (feminine) with marked number (plural). The gender on Num will then be transferred to D, determining the gender of the whole DP, on the assumption that the higher gender always wins (Alexiadou 2004; Steriopolo and Wiltschko 2010; Matushansky 2013). The different types of Num in TA are presented in (64).

(64) **Types of Num in TA**

- a. Num [-PL] singular nouns
- b. Num [+PL] plural animate nouns
- c. Num [+FEM,+PL] plural inanimate noun

Num is unspecified for gender in (64-a) and (64-b), since the conditions for gender neutralizing on *n* are not present: this only happens for Num heads marked [+PL] in the context of an uninterpretable gender on *n*, as in (64-c).

In the analysis presented above, I argued that a gender switch in the plural of

TA inanimate nouns is because the Num head that is specified [+PLURAL] bears a [+FEMININE] feature that can only override uninterpretable gender features on *n*. This analysis is in line with other analyses of gender viewed as an inherent feature of number affixes (Farkas, 1990; Bernstein, 1993; Ritter, 1993; Giurgea, 2008; Croitor and Giurgea, 2009). Like on the *n* head, I assume that gender features on Num can be interpretable or uninterpretable. This becomes even more relevant in the discussion about the interpretability of the feminine marker in singulative nouns in [Chapter 3](#). In the case of inanimate plurals in TA, the feminine gender specified on Num has no semantic contribution. Therefore, it goes without saying that gender is an uninterpretable feature on Num in this case.

2.5.3 Alternative approaches

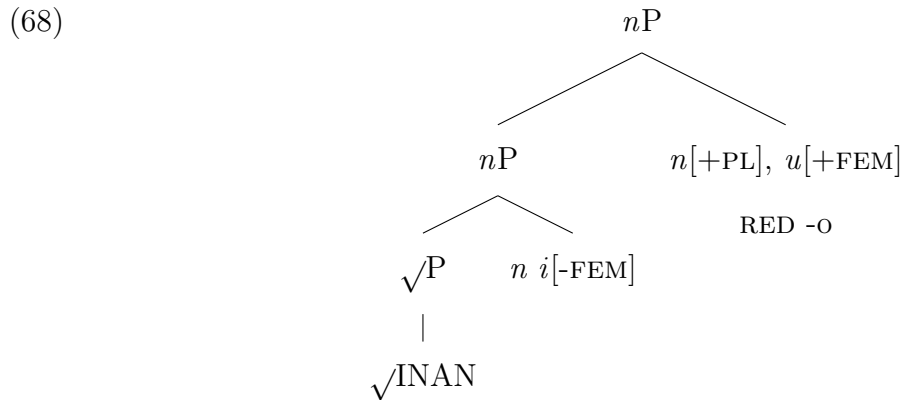
The analysis presented above does not provide a single locus for gender, as does the one offered by Kramer (2009, 2015). A unified analysis of gender locus would have been simple and elegant, and thus desirable for the theory, but it would not have accounted for the data at hand. Certain analyses account for gender-number dependencies by arguing that the plural can be a feature on *n*. This is notably the case for Lecarme’s analysis of gender polarity in Somali. Somali nouns can switch gender in the plural (65)-(66).

- | | | | | | | |
|------|----|--|------|----|--|----------|
| (65) | a. | díbi-ga
bull-DEF.MASC
'the bull (male)' | (66) | a. | náág-ta
woman-DEF.FEM
'the woman (female)' | |
| | b. | dibí-da
bull.PL-DEF.FEM
'The bulls (female)' | | b. | naag-á-ha
woman-PL-DEF.MASC
'the women (male)' | [Somali] |
- (Lecarme, 2002)

Based on the data in (66)-(65) and assuming that plural suffixes can bear gender specification in Somali, [Lecarme \(2002\)](#) argues that these plural suffixes behave like derivational suffixes in productive (category-changing) derivational morphology. Many properties of plural formation in Somali provide evidence for the idea that plural formation is a derivational process (hence, on *n*). First, the choice of a plural strategy is root-specific. A particular plural suffix has a contextual feature to limit its attachment to certain roots. Lecarme claims that this kind of arbitrary relationship is more characteristic of the relationship between *n* and a root, than it would be for an inflectional head like Num.

Second, plural suffixes in Somali have some selectional restrictions. Plural affixes never attach to categories other than nouns, and they select particular stems. In other words, the affixation rule must already “know” whether the root is monosyllabic or bisyllabic, and whether it is masculine or feminine.

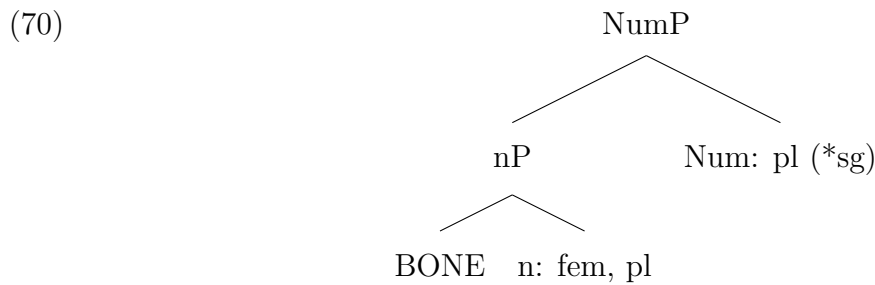
A third piece of evidence that the plural is derivational in Somali is that there can be multiple plural forms for the same noun, that is, plural marking is non-deterministic for some roots in Somali. This is especially the case for nouns that can form prosodic plurals and also have one or two other plural forms involving infixation



Acquaviva (2008, 2009) also argues that *n* can carry number information. He claims that this is notably the case for Italian irregular plurals such as *ossa* ‘bones as connected parts, not self-standing pieces’. Consider the example in (69), keeping in mind that the “regular” plural affix for a masculine noun is *-i*, and that *-a* is usually associated with feminine singular nouns.

- (69) a. *osso* ‘bone’ [[BONE] n: masc]
 b. *ossa* ‘bones’ [[BONE] n: fem, pl]

The resulting structure would, thus, be the following:



Acquaviva notes that according to this analysis, feminine plural forms such as *ossa*

are not feminine nouns in a plural context – nor masculines that change gender in the plural like in Romanian – but inherently plural feminine nouns. They are root nominalizations that induce a distinctive interpretation that is different from the one obtained from the masculine plural alternant, if applicable. In the case of *ossa*, for example, the regular form *ossi* (masculine plural) is available, and has the regular meaning ‘bones’, not necessarily as ‘connected parts’, as in *ossa*.

While analyzing irregular plurals as being on *n* is a desirable alternative that keeps all gender features on the same head, it is not a viable option for the sound plurals of TA. It is true that the pluralization strategy in Arabic depends on the shape of the root, however, there are only two available strategies, namely the sound (affixational) and the broken (stem internal) plurals. Both are very regular in their morphology and in their meaning. The sound plural, for instance, only comes in two shapes, one for the masculine and one for the feminine. Gender change is only attested with inanimate nouns. The broken plural, which is discussed in [Chapter 4](#), follows a regular prosodic pattern and is very predictable from the shape of the singular for a native speaker of Arabic. In addition, none of these plural strategies are associated with special meaning, as is the case for *ossa*. Therefore, I assume that the plural in Arabic is always associated with the Num head.

2.5.4 Summary

In this section, I have addressed the case of TA inanimates, which always take the feminine suffix *-at* in the plural. I have established that they are feminine and have argued that the feminine is a feature on number. Gender features on Num can

only override uninterpretable features on *n*, since they are weaker than interpretable gender features. While Ritter (1993) argues that a language can have gender features on *n* or on Num, I claim that TA is a language where gender is both on *n* and Num, by presenting the language-specific rules that determine the gender of the DP in TA.

2.6 Conclusion

In this chapter, I described the data pertaining to gender assignment in the TA count system and discussed the existing definitions of gender, its functions, and possible loci in the DP. I argued that gender has a nominalizing role and is hosted on *n* in TA, following previous authors (Kihm 2005; Lowenstamm 2008, 2012; Kramer 2009, 2014, 2015, among others). I used evidence from root and pattern morphology to provide stronger support for the idea that noun formation is done outside the root, with gender being part of the nominalization operation.

I adopted Kramer's (2009; 2014; 2015) analysis of gender in Amharic, with gender features of different interpretability values on *n*, depending on whether or not they bring a semantic contribution to the meaning of the *n*P. I showed that animate nouns have interpretable gender on *n*, while inanimate nouns have uninterpretable gender on *n*.

Finally, I addressed the issue of gender marking in inanimate plurals. Following previous Num-based accounts of gender (Farkas, 1990; Ritter, 1993; Bernstein, 1993; Giurgea, 2008; Croitor and Giurgea, 2009), I argued that feminine plural markers of inanimate nouns were the result of the Num head being marked with an unin-

interpretable [+FEM] marker. In this analysis, I account for the principle of animacy hierarchy by claiming that the semantic gender of an animate noun is more likely to be preserved across numbers than arbitrary gender of inanimate nouns.

In the next chapter, I discuss collective nouns in TA, that constitute a system parallel to that of count nouns as described in the present chapter. Gender allotment is based on number in the collective class. I discuss the gender features of collective nouns in terms of their interpretability and structural locus.

Chapter 3

The singulative

3.1 Introduction

This chapter examines the second number system of TA, the “collective system”. Recall that the number systems are distinguished in terms of the way they mark number. In the “count system”, discussed in [Chapter 2](#), the singular is morphologically unmarked, and the plural is marked, either by suffixation (sound plural) or stem change (broken plural). In the collective system, sum-referring nouns¹ are morphologically unmarked, and individual-referring nouns are marked by the suffix *-a*. The coexistence of two parallel number systems raises many questions, namely what is the basis of classification of each noun? How does a speaker learn which number system to use with an established noun, a neologism or a borrowing? This boils down to determining what properties (semantic or other), if any, characterize

¹I am using “sums-referring” as a pre-theoretical term. I examine the idea that collectives are number neutral, as suggested by [Ojeda \(1992\)](#); [Zabbal \(2002\)](#); [Fassi Fehri \(2012\)](#); [Mathieu \(2013\)](#). I further develop this analysis in [Chapter 4](#).

entities or nouns belonging to the collective system. Furthermore, the question of the role of gender is crucial in this system, especially following the discussion in [Chapter 2](#), about gender being a nominalizer in the count system of TA. In the case of collective nouns, it seems that gender plays a different role, one that is linked with number interpretation. One of the goals of this chapter is to discuss this issue and propose a way to formalize the role of gender in the collective system of TA. I start by defining the properties of the collective class and establishing the characteristics of the nouns that make up this class in [Section 3.2](#). I conclude that, although there is some sort of semantic coherence within the nouns of the collective class, the true motivation behind a noun's classification into one class or another is the way the language chooses to morphologize this particular noun. Therefore, I argue that the collective is a class feature marked on n , and not an inherent characteristic of the root. Based on my definition of the collective class and on the existing analyses in the literature, I propose, in [Section 3.4](#), that the singulative is an inflectional operation, and that in TA, it translates into an interpretable gender feature on Num. This echoes my treatment of feminine plural animates in TA, presented in [Chapter 2](#), that I justified by the presence of an uninterpretable number feature on Num. A Num-based analysis of the singulative has, among others, the advantage of taking into accounts singulative systems that do not rely on gender marking. In [Section 3.3](#), I examine alternative analyses of the singulative, focusing on cases where it is marked by a gender feature. I pay close attention to how and where the features are hosted in each analysis and discuss the advantages and limitations of each, while commenting on their potential predictions. The conclusion is presented in [Section 3.5](#), where I

summarize the claims presented in the chapter about the nature of collective nouns and the singulative operation.

3.2 The collective system: definition and data

Crosslinguistically, the plural is most often morphologically marked as opposed to the singular (Greenberg, 1972). This can be seen in (1) for English. The singular form is the basic form and surfaces with no added suffix (1-a), while the plural form is obtained by adding the suffix *-s* to the word *dog*, as shown in (1-b).

- (1) a. one dog
b. two dog-*s*

In other languages including Arabic, there is a system parallel to the one described above, where the unmarked form, the *collective*, is used to refer to sums despite the absence of plural morphology. The collective form can derive a singular form traditionally called the *singulative*. Arabic grammarians refer to the singulative as *esm el waḥda* ‘noun of unity’ (Wright 1933, 147, Ojeda 1992). In the following sections, I explore the different instances of the singulative across languages and discuss the characteristics defining the class of nouns on which the singulative can operate.

3.2.1 Singulative marking across languages

While in a number of languages such as Arabic, the singulative is expressed by gender shift, in other languages, the singulative is marked by a dedicated suffix that does not involve gender (Dimmendaal 1983, 2000; Hieda 2006 for Nilotic languages; Grimm 2012 for Niger-Congo). (2) shows examples of the singulative expressed through gender, and (3) shows examples of the singulative expressed through dedicated suffixes.

- (2) a. ħmem ħmem-a [Standard Arabic (Wright 1933: 147)]
pigeon.MASC.COLL pigeon-FEM.SING
'pigeons, a pigeon'
- b. geot geot-enn [Breton (Trépos 1956: 67)]
grass.MASC.COLL grass-FEM.SING
'grass, a blade of grass'
- c. ħobz ħobz-a [Maltese (Weiss 1845: 7)]
bread.MASC.COLL bread-FEM.SING
'bread, a loaf of bread'
- (3) a. gwenyn gwenyn-en [Welsh (King 2015: 75)]
bee.FEM.COLL bee-FEM.SING
'bees, a bee'
- b. abwyd abwyd-yn [Welsh (King 2015: 77)]
worm.masc.coll worm-MASC.FEM
'worms, a worm'
- c. mùó mùó-rùò [Dagaare (Grimm, 2012)]
grass.COLL grass-SING
'grass, a blade of grass'
- d. àyó àyó-k [Aiki (Runga) ((Nougayrol, 1990, 47-48)]
leaf.COLL leaf-SING
'leaves, a leaf'

In the examples in (2), the singulative affix also marks a gender shift from masculine to feminine. In (3), however, we find examples of dedicated singulative suffixes that mark no other feature. Note that in Welsh, the singulative affix can vary according to the base gender: feminine collectives take the suffix *-en* (3-a), and masculine collectives (3-b) take the suffix *-yn*. While the singulative is marked for gender, that gender does not vary from the one found on the collective base and hence we cannot say that the number distinctions are marked by gender in this language. In Dagaare (3-c), the singulative suffix is an allomorph of the plural marker. When affixed to certain bases, the suffix marks the plural, and when affixed to other types of bases, the suffix marks the singulative operation. For Grimm (2012), the interpretation of the suffix depends on the propensity of a noun to be inherently plural. For inherently plural nouns, the plural tends to be morphologically unmarked (in line with the semantics) and, therefore, the morpheme *-ri* (or its allomorph *-ruu*) marks the singular interpretation of the nouns. For inherently singular nouns, the morphological base is singular (again in line with semantic markedness), and the plural is marked by the suffix *-ri*. In other words, in the system of Dagaare, the morpheme *-ri* is a suffix for *marked* number.

3.2.2 Definitions and properties of collective nouns

Focusing on Arabic, I now review some well-known definitions of the singulative and the properties of nouns belonging to the collective class. In Wright's (1933) grammar, the *Nomina Unitatis*, or nouns of individuality, designate one individual out of a genus, or one part of a whole that consists of very similar parts. Ojeda (1992)

defines Arabic collective nouns as basic lexical items that indicate either “a substance or material in the mass” or “a collection of objects viewed as a totality without reference to the individual members” (referring to [Erwin 2004](#), 166). Singulative nouns are lexical items that are derived from collectives and refer either to “a specific quantity of the substance” or to “an individual member of the collection”. [Greenberg \(1972\)](#) refers to languages with a collective system as “three term systems” where a collective which cannot be used with numerals is opposed to a singulative with its own singular and plural (the plural of the singulative is discussed in [Chapter 4](#)). [Greenberg](#) stresses the obvious analogy between mass nouns and collectives, based on their similar syntactic distributions. He also notes that in languages with a collective system, the collective tends to be used in some instances for nouns designating materials and even liquids, in which case the singulative designates quasi-units (“quasi” in the sense that they require a unit of measure).

The distribution of collective nouns is similar to that of mass nouns: they are syntactically singular but cannot be pluralized (4-b) ((4-a) is a count noun, for purposes of comparison),² they do not combine directly with numerals (5)³ and they can combine with cumulative quantifiers (6).

²Very few exceptions exist, where the collective can be directly pluralized, resulting in a very specialized reading, e.g. ‘several types of x’ or ‘an abundance of x’, see ([Acquaviva, 2008](#), 47). This operation is not productive in TA.

³Combination with numerals higher than *one* also depends on pluralization. The best way to test this type of construction would be by using the numeral *one*, however, singular nouns are used bare in Arabic and cannot combine with a numeral. For example, **wehed qattus* ‘one cat’ is ungrammatical in TA. The correct construction would be *qattus* ‘one/a cat’. Therefore, I am aware that combination is not possible with collective nouns due to the fact that they cannot be directly pluralized. However, non combination with numerals indicates that, although a collective noun refers to sums, it has a syntactic status that is different from plurals, which also refer to sums but can combine with numerals.

- (4) a. qattus → qtates [Tunisian Arabic]
 cat.MASC.SG cat.PL
 ‘a cat, cats’
- b. tut → *atwat
 berry berry.PL
 ‘berries, (intended) berries’
- (5) a. xamsa qtates [Tunisian Arabic]
 five cat.PL
 ‘five cats’
- b. *xamsa tut
 five berry.MASC.COLL
 ‘(intended) five berries’
- (6) a. *barfa qattus [Tunisian Arabic]
 lot cat.MASC.SG
 ‘(intended) many cats’
- b. barfa tut
 lot berry.MASC.COLL
 ‘a lot of berries’.

In what follows, I discuss the characteristics of collective nouns in terms of cumulativeness (Quine, 1960) and divisiveness (Cheng, 1973; Krifka, 1989). Cumulativeness is the property concerned with sums and is defined as follows:

- (7) A noun is cumulative iff it denotes a cumulative predicate.
 A predicate p is cumulative iff any sum of parts that are p are also p .
 (Deal, 2017)

Water, for instance, is cumulative, because if a is water and b is water, then $a + b$ is

water. *Cat*, on the other hand, is not cumulative, because if a is a cat and b is a cat, then $a + b$ is not a cat. $\llbracket \textit{cat} \rrbracket$ is a predicate holding only of individual cats, but not groups or pluralities thereof. Therefore, *cat* is not cumulative. This type of inference correlates with morphosyntactic differences between cumulative and non-cumulative nouns, i.e., their ability to combine with cumulative predicates and the fact that they cannot be pluralized.

Based on Chierchia's (1998) analysis, collective nouns are no different from mass nouns, since they both pattern semantically like plurals (they are both cumulative). Chierchia makes no distinction between collective and object-mass nouns, such as *furniture* and *jewelry*, claiming that both types of nouns have the same syntactic behaviour due to their inherent plurality. This led Zabbal (2002) to explore the similarities between Arabic collectives and object-mass nouns (dubbed "*furniture*-type nouns" in his work). Zabbal concludes that *furniture*-type nouns and Arabic collectives, in addition to being both cumulative, possess the same property called "individuation". This is the second countability distinction (also referred to as "divisiveness"), which is concerned with parts (Cheng, 1973; Krifka, 1989) and is defined as follows:

(8) A noun is divisive iff it denotes a divisive predicate.

A predicate p is divisive iff any part of something that is p is also p .

(Deal, 2017)

A noun is *individuated* if it does not refer to a homogeneous whole or in other words, if it refers to a number of distinct entities that the semantics can tell apart.

Zabбал described two ways of distinguishing entities as a whole and uses them both as diagnostic tools for the property of individuation. The first one is the use of spatial relations like the predicates *arrange*, *among*, or *between*. These predicates require a semantically plural argument whose “parts” are spatially distinct. Examples are provided in (9) for cumulative and individuated nouns like *dogs*, *furniture*, and *cattle*.

- (9) a. I stood among the dogs/*the dog/*the committee/*the water.
b. I zigzagged between the furniture/*the dog/*the committee/*the water.
c. I arranged the cattle/*the dog/*the committee/*the water.

In (9-a), the predicate *among* is predicated of *the dogs*, which denotes a sum of atoms (a plurality). Each dog is one of these atoms. The adverbial PP *among the dogs* describes a location. The spatial meaning associated with the predicate *among* requires that this location be inside an area populated by the dogs. Thus, the individual dogs must be spatially separated. Similarly, in (9-b), the act of zigzagging between the furniture requires that a space exist between the pieces of furniture. Finally, in (9-c), *arrange* literally means to organize spatially in some pattern. Again, the things being arranged, in this case the head of cattle, must occupy spatially distinct coordinates. These predicates cannot be predicated of NPs headed by non-individuated nouns, such as *the dog* (singular count noun), *the committee* (group noun), and *the water* (substance/mass noun), as illustrated in (9).

The second way to distinguish individual entities is to count them out, to label them, or to number them (sequentially). For example, the act of numbering a group of entities has the desired effect, namely to set each entity apart from the others

by giving it a unique number. Examples are provided in (10), for NPs headed by individuated nouns, such as *the dogs*, *the furniture*, and *the cattle*.

- (10) a. I counted out the dogs/#the dog/#the committee/#the water.
b. I labelled the furniture/#the dog/#the committee/#the water.
c. I numbered the cattle/#the dog/#the committee/#the water.

Predicates, like *count out*, *label*, *list*, are underlyingly indexical relations, where a part is indexed with respect to the other parts of a whole. If an entity is individuated, then its parts can be counted out, labelled, or numbered. So, in (10-a), the individual dogs are counted out; in (10-b), the individual pieces of furniture are labelled; and in (10-c), the individual heads of cattle are numbered. The arguments in (10), namely *the dogs*, *the furniture*, and *the cattle*, refer to the complete set of entities being indexed. The individuation tests reveal that cumulativity does not suffice to describe a noun in terms of countability. Some cumulative nouns are individuated, and some are not. Both countability distinctions are necessary in order to get a full picture of the noun's denotation (see also Deal 2017 for a non-binary treatment of the mass-count distinction).

Applying this test to Arabic collectives, it can be observed that a collective can also be an argument of an individuating predicate. The data in (11) illustrate this.

- (11) ?inna-ni ?a-jmaʕ-a n-naml-u. [Lebanese Arabic]
PART-1.SG 1-counted-1 the-ant.COLL-NOM
'I counted the (individual) ants.'

(Zabbal, 2002, 98)

✓ In a video game where the player has to catch and eat as many fish as possible: Myriam’s character ate five sardines and Sarah’s ate one salmon.

x In a video game where the player has to catch and eat as many fish as possible: Myriam’s character ate one salmon and Sarah’s ate five sardines.

The example in (12-b) shows that a collective noun like *ḥut* ‘fish’ combines with distributive adjectives, therefore it is expected to denote an individuated semi-lattice. However, the comparative constructions tests (13) bring out another dimension of collective nouns, that the combination with distributive adjective did not show. While the former test suggests that collective nouns are individuated, the latter shows that their denotation is ambiguous between individuated and unindividuated semi-lattices. If this prediction is correct, then we should expect distributive adjectives to combine only with individuated collectives. This prediction is borne out in (14).

- (14) Mariem kle-t akthar ḥut sghir men Sarah. [TA]
Myriam ate-3FEM.SG more fish.MASC.COLL small.SG than Sarah
‘Myriam ate more small fish than Sarah.’

✓ In a restaurant: Myriam ate five small fish and Sarah ate one small fish.

x In a restaurant: Both Myriam and Sarah have a small fish on their plates.

Myriam ate a bigger portion of her fish than Sarah did.

(*Small* cannot refer to a portion in this context.)

Unlike Zabbal’s conclusion, the tests in (13) and (14) suggest that Arabic collective nouns have an ambiguous meaning that can vary in terms of individuation. The term *ḥut* ‘fish’ can refer to individual fish or to substance or “flesh”. This is also the

case for its English and French equivalents *fish* and *poisson*, as is shown in (15) and (16) with the use of indefinite determiners versus partitive constructions.

- (15) a. A fish
b. Some fish (ambiguous between indefinite individual and partitive)
- (16) a. Un poisson [French]
INDEF fish
'a fish'
b. Du poisson
PART fish
'Some fish (partitive)'

In the same vein, [Ojeda \(1992\)](#) also notes that some collective nouns such as *djaaj* 'chicken' can be viewed as a kind of food or as a species.

This brings us to the domain of English nouns like *furniture*, *jewelry*, and *footwear*, which are often referred to as "object-mass" ([Bale and Barner, 2009](#)), aggregates ([Deal, 2017](#)), or "furniture-type" nouns ([Zabbal, 2002](#)). The nouns in this category have a mass noun distribution. They cannot be pluralized (17), cannot combine with numerals (18) and are cumulative (19).

- (17) *the furnitures
(18) *one footwear
(19) How much footwear did you buy?

In terms of the parts-based countability distinction, *furniture*-type nouns turn

out to be unambiguously individuated based on all tests, as shown in (20).

- (20) a. I arranged the furniture/jewelry.
b. I counted out the furniture/jewelry.
c. I bought small furniture/jewelry.
d. I bought more furniture/jewelry.

(The comparison is based on number of pieces of furniture/jewelry, and not in terms of mass or volume).

Unlike Arabic collective nouns, *furniture*-type nouns are unambiguously individuated. Zabbal (2002) compares the denotation of collective nouns to that of *furniture*-type nouns. However, based on the results yielded from the comparative construction tests, I propose that this is not exactly the case: object-mass and collective nouns are distinguished by the individuation property. Collective nouns are more “massy”; they are ambiguous between individuated and unindividuated reference, unlike object-mass which are unambiguously individuated. In the next section, I discuss the differences between Arabic collectives and other types of nouns that are similar in some aspects, including object-mass nouns. I show that collective nouns, despite their resemblances to other types of nouns, have unique properties and are distinguished by the fact that they belong to a different morphological class.

3.2.3 What the collective is not

The Arabic collective can be confused with all sorts of other nouns for various reasons. For example, one may be tempted to compare collectives with group nouns (e.g., *committee*), object-mass nouns (e.g., *furniture*), or Arabic singular nouns referring to ethnic groups or members of a profession (dubbed *pluratives* by [Fassi Fehri 2012, 2018a,b](#)). All these comparisons are due to resemblances with collective nouns in certain specific characteristics. However, I argue that the collective is a distinct morphological class where there is an inverse morphological marking between the sums-denoting and the individual-denoting nouns, an essential characteristic that all other types of nouns do not have. I now review each one of these categories and show, through relevant tests, that they all differ from the collective, and conclude with a more fine-grained definition of the collective noun.

The first type of noun that one may confuse with a collective noun is the object-mass noun. As discussed above, [Zabbal \(2002\)](#) compared Arabic collectives to *furniture*-type nouns (or object-mass), lending them similar denotations. Both types of nouns have mass (cumulative, no pluralization and no combination with numerals) distribution and are individuated (combine with individuated predicates). On the basis of the results obtained from the comparative constructions test, Arabic collectives do not always refer to individuated entities, but are ambiguous between substance and individuated reference. Object-mass nouns, on the other hand, are always individuated. In addition to this difference, collectives represent a special morphological category: they are always masculine and can derive singulatives through gender shift. This is not the case for object-mass nouns. Although object-mass nouns can be the

semantic equivalents of collectives, they are not part of the same morphological class. Object-mass nouns in TA, for example, can be of any gender (masculine or feminine), which makes the singulativization through feminine gender impossible. This is illustrated in (21) with the TA object-mass nouns *siegha* ‘jewelry’ and *mobilia* ‘furniture’. These nouns behave like English object-mass nouns in terms of their sums-based and parts-based properties: they are cumulative (21-a), they cannot be pluralized (21-b), they do not combine with numerals (21-c), and they are individuated (21-d).

- (21) a. el siegha/mobilia el kol [Tunisian Arabic]
the jewelry/furniture the all
‘All the jewelry/furniture’
- b. *el siegha-at/mobilia-at
the jewelry-PL/furniture-PL
‘(intended) the jewelries/furnitures’
- c. *xamsa siegha/mobilia
five jewelry/furniture
‘(intended) five jewelries/furnitures’
- d. hsebt el siegha/mobilia.
counted.1.SG the jewelry/furniture.
‘I counted (out) the jewelry/furniture.’

Despite these characteristics, which make them very similar to collectives, object-mass nouns like *siegha* ‘jewelry’ and *mobilia* ‘furniture’ are not part of the same class. First, their collective counterparts are feminine,⁴ and not masculine like all collective nouns (22). I insist that in this case, the feminine noun does *not* refer to

⁴This is accidental, there are also instances of masculine object-mass nouns in Arabic. What I want to show here is that gender is arbitrary in this type of nouns, and not contingent on number as it is the case in the collective system.

the singulative, but to sums. Furthermore, and this naturally follows from the fact that they are not always masculine, TA object-mass nouns do not derive singulatives via gender shift (23-a). The only way to obtain an individual reading out of an object-mass noun is by combining the noun with a ‘portion of’ predicate (22-b) as in English (e.g., *blade of grass*, *head of cattle*, *sheet of paper*). This type of predicate is not a classifier nor is it part of an inverse system since individuation is not associated with a specific morpheme, as it is the case for the singulative in Arabic [Greenberg \(1972\)](#).

- (22) Lqit siegha masrouq-a. [Tunisian Arabic]
 found.I jewelry stolen-FEM.SG
 ‘I found stolen jewelry.’

✓I found a bag with a necklace, earrings and a ring.

? I found a bag with a ring.

- (23) a. siegha/mobilia → *siegha-a/mobilia-a [Tunisian Arabic]
 jewelery/furniture jewelry-FEM.SING/furniture-FEM.SING
 ‘jewelry, (intended) a piece of jewelry/furniture’

- b. qatfa siegha/mobilia
 piece.SG jewelry/furniture
 ‘a piece of jewelry/furniture’

Given that the morphologically marked singulative is the hallmark of the collective system, it follows that object-mass nouns are of a different kind. They rather seem to be accidental nouns referring to a collection of objects that are of similar use.

There is a second kind of syntactically singular noun with a group interpretation

that, I argue, is distinct from the collective. The “plurative”⁵ (Fassi Fehri, 1984, 1988, 2004, 2012, 2018a,b) is an nP designating a group unit, which is morphologically characterized by the feminine suffix *-a* and which triggers feminine singular agreement on targets (verb or adjective). The result is usually interpreted as an integrated whole, as opposed to a distributive interpretation. Some well-known plurative patterns include expressions of professional groups or corps as illustrated in (25), and nPs referring to ethnic or regional groups (26).⁶

- (25) a. *saydal-ii* ‘pharmacist-masc’ → *sayaadil-a(t)* ‘the corps of pharmacists’
 [Standard Arabic (Fassi Fehri 2018b, 10)]
- b. *xayyat* ‘tailor’ → *xayyat-a* ‘tailor-fem’, ‘tailors’ (the corps of tailors)
 [Moroccan Arabic (Fassi Fehri 2018b, 10)]
- (26) a. *?afriq-ii* ‘african-masc’ → *?afaariq-a(t)* ‘Africans’
 [Standard Arabic (Fassi Fehri 2018b, 10)]
- b. *jubl-ii* ‘inhabitant of the mountain-masc’ → *jbal-a* ‘inhabitants of the mountain’
 [Moroccan Arabic (Fassi Fehri 2018b, 11)]

⁵Fassi Fehri borrows the term “plurative” from Africanist literature (Dimmendaal, 2000; Mous et al., 2008; Blench, 2013; Treis, 2014). Basically, the plurative is the opposite process to the singulative, where instead of deriving a singular form from the basic plural, the process takes the opposite path by deriving a plurative from a more basic singular (the collective), as illustrated in (24).

- (24) ‘collective’ → singulative
 plurative ← ‘collective’

⁶According to (Fassi Fehri, 2018b, 11), morphologically, it is not important that pluratives (most often) use broken plural patterns as their base as it is the case in the examples from Standard Arabic. It is, however, essential that they are gendered by the suffix *-a(t)* and most importantly, trigger feminine singular agreement on targets.

mittee are considered to be hybrid nouns carrying mismatching syntactic/semantic ϕ -features (Landau 2015; Corbett 2000, 2015; den Dikken 2001; Wechsler and Zlatic 2003; Danon 2011, 2013; Smith 2015). Both sets of features can be accessed, which explains the variation observed in certain English dialects, as illustrated in (29).

- (28) El xayyat-a t-exdem/y-exdm-u bekri. [Tunisian Arabic]
 the tailor-FEM.SG FEM.SG-work/MASC-work.PL early
 ‘The tailors work early.’
- (29) a. The committee is meeting next week.
 b. The committee are meeting next week.

Collective nouns, on the other hand, do not share these properties. They only permit masculine singular agreement (30), which distinguishes them from pluratives and other group nouns.

- (30) El nemmel y-ejri/*y-ejri-w. [Tunisian Arabic]
 the ant.MASC.COLL MASC-run.SG/MASC-run.PL
 ‘The ants are running.’

Moreover, unlike collective nouns, pluratives cannot derive singulative nouns (this would be impossible since they are already feminine) (31), and consequently cannot be considered as part of the collective class.

- (31) xayyat-a → *xayyat-a-a [Tunisian Arabic]
 tailor-FEM.SG tailor-FEM.SG-FEM.SING
 ‘tailors, (intended) a tailor’

follows that a group noun is an atom ([Zabbal, 2002](#), 61-62). This means that group nouns are count nouns, and not mass. This is further illustrated in (35), where we can see that unlike Arabic collectives, group nouns can be pluralized (35-a) and combined with numerals (35-b).

- (35) a. The committees are meeting tomorrow.
b. Four committees are meeting tomorrow.

(35-a) is an instance of a pluralized group noun and means that several committees are meeting. (35-b) shows an instance of a group noun combined with a numeral, meaning that four different groups, and not members thereof are meeting.

Along the same lines, [Barker \(1992\)](#) argues that groups are of dual nature and can denote either atoms/individuals with no internal part structure or sets that are at least partially determined by the properties of their members. The group as an atom differs, semantically and syntactically, from both plurals and conjunctions, contrary to the widespread view (in [Bennett 1974](#); [Link 2002](#); [Landman 1989](#)). This analysis is confirmed by:

- a. uses of names of groups as rigid designators;
- b. parallels between group nouns and measure nouns;
- c. the distribution of groups in dialects of English.

Among the salient syntactic properties of atomic group nouns are the following:

- a. Groups can be pluralized (*committees, armies*);
- b. They can be counted out (*two committees*);
- c. They can take an *of* phrase containing a plural complement (*an army of children, *an army of a child*).

Based on these observations, we can conclude that group nouns are count nouns with a sums denotation, and not mass nouns. Based on supporting evidence from syntactic distribution, I have established that Arabic collective nouns behave like mass nouns. Therefore, collectives are not group nouns.

3.2.4 The collective is a class type on n

To wrap up this section, which aims to define the collective class, I now return to the different definitions of the collective in the existing literature and compare them with the observations and results drawn from the tests above. Collective nouns have often been described as nouns denoting “inherently plural objects”, “a collection of objects as a totality” or “masses or substances” (Erwin, 2004, 166). They often refer to small entities, vegetation, fruits, nuts, small animals, insects (Gaudefroy-Demombynes and Blachère, 1952). Along the same lines, Talmoudi (1980, 132) views the collective as denoting either “a collection of things or animals regarded as a unit”, or else “a mass or volume”. Wright (1933, 147) describes the collective as expressing “the genus or whole”, and Abdel-Massih et al. (1981, 49) define the collective as a noun that designates “a class or mass of like things without counting

the units that make up the mass”. For [Fleisch \(1961, 65\)](#), a collective noun is “the mass wherein the individuality of the ‘amassed’ is effaced”. All of these definitions share the idea of inherent plurality and objects viewed as substances. The general idea is that the morphological marking of these systems aligns with the fact that singularity is semantically marked in these nouns. Along these lines, [Grimm \(2012\)](#) investigated the correlation between number markedness and the semantic level of individuation of a noun in the inverse system of Dagaare (Gur; Niger-Congo). His hypothesis is that individuated nouns (semantically unmarked in the singular) are more likely to be morphologically marked in the plural. On the other hand, non-individuated/inherently plural nouns (semantically unmarked in the plural) tend to be morphologically marked in the singular. His predictions are borne out. According to his analysis and as shown in example (3-c) of this chapter, the suffix *-ri* signals the semantically marked number in Dagaare. In the case of substance and inherently plural nouns, *-ri* marks the singulative. In this kind of system, morphological markedness is aligned with semantic markedness. This view aligns with the definitions of collective nouns provided above, where it is a question of “things viewed as a totality” and collections of things regarded as “a whole or unit”. This also explains the ambiguous individuation readings as shown in (13), where a noun of the collective class like *ħut* ‘fish’ can denote both a substance and a collection of animals. There is undoubtedly a tendency for collective nouns to denote substances or small things that are not distinguishable. However, this property does not suffice to classify nouns in the collective class. For instance, different dialects have different ways of classifying nouns; one and the same entity can have a collective noun in one

dialect and a count noun in another. This is the case for the noun *baqar* ‘cows’.⁷ Although this noun is sums-referring in all dialects, it is a broken plural triggering plural agreement in TA (36), and a singular collective in other dialects such as Lebanese Arabic (37) (Zabbal, 2002).

- (36) a. xamsa bgar harb-u. [Tunisian Arabic]
 five cow.PL escape.-3MASC.PL
 “Five cows ran away.”
- b. haaḏa baqar [Classical Arabic (Abd-Rabbo 1990)]
 this.MASC.SG cow.MASC.SG
 “these cows”
- c. baqar baqar-a [Lebanese Arabic (Zabbal 2002)]
 cow.MASC.COLL cow-FEM.SING
 “cows, a cow”

The example in (36-a) shows that the TA noun *bgar* ‘cows’ triggers plural agreement and, unlike collective nouns, can also combine with the numeral *xamsa* ‘five’. We can therefore consider *bgar* to be the broken plural form of the noun *bagra* ‘cow’. In Classical and Lebanese Arabic, however, it appears that *baqar* ‘cows’ triggers masculine singular agreement (36-b-c). In addition, Abd-Rabbo (1990) emphasizes that the expression **xamsu baqar* ‘five cows’ is ungrammatical in Classical Arabic. This confirms that *baqar* is part of the collective system in both Classical Arabic and the Lebanese dialect.⁸ This divide also exists in substance nouns. For example, the

⁷Referring to (Corbett, 1996, 111-14), Acquaviva (2008) suggests an alternative view of such examples, where syntactically, collectives that are high in the Animacy Hierarchy can trigger plural agreement, as an instance of semantic-based agreement. However, since the agreement being always plural in TA, I argue that *bgar* is not part of the collective class in TA.

⁸This confusion is unsurprising, since broken plurals are stem-internal operations and nothing in their shape distinguishes them from collective nouns (Acquaviva, 2008). Their true difference

word for ‘mud’ is *tabʕa* in TA and *teen* in Saudi Arabic. While *tabʕa* is a substance-referring noun with no singulative counterpart (37-a), *teen* has a singulative, *teen-a*, which refers to a chunk of mud (37-b).

- (37) a. *tabʕa* → **tabʕa-a* [Tunisian Arabic]
 mud mud-FEM.SING
 ‘mud, (intended) a chunk of mud’
- b. *teen* → *teen-a* [Saudi Arabic (Mathieu, 2014)]
 mud.MASC.COLL mud-FEM.SING
 ‘mud, a chunk of mud’

Although there is a consensus that substance and inherently plural nouns tend to be part of the collective system, the variation in (36) and (37) shows that the count/collective classification is not encoded in the root denotation of a noun, but is rather part of its grammatical information. Therefore, I argue that whether a noun belongs to the count or collective class, it is encoded in *n*, as opposed to being specified in the lexicon. This yields the nominal representation in (38), where the root combines with a *n* marked [+COLL]. The locus of gender and how it is interpreted in this class is discussed in [Section 3.4](#).



resides in their triggered agreement patterns and whether or not their singular counterparts are morphologically marked.

To sum up, this section defined the collective class and the singulative. I discussed some well-known definitions of the Arabic collective. Most of the definitions revolved around the central theme of “massiness” and inherent plurality. I then tested the claims drawn from these definitions by testing the distribution of collective nouns to define them in terms of countability distinctions (both sums and part-based). It was established from these tests that Arabic collectives are cumulative and ambiguous in terms of individuation. This distinguishes them from object-mass nouns, which are known for being individuated mass nouns. This corroborates the fact that most definitions of the collective compare them to substance nouns. I also showed that the Arabic collective is a class that is distinct from pluratives and group nouns. This conclusion was based on their distribution and, more importantly, on the fact that the latter types of nouns cannot morphologically derive singulative counterparts. Finally, I established through cross-dialectal variation in noun classification that a noun does not come from the lexicon specified as [+/-COLL], but that this feature is instead part of the functional structure, on the *n* head, more specifically.

3.2.5 Should we call it gender?

Based on its interpretation, the singulative appears to be closer to number than gender. It operates on nouns that are already nominalized, and although marked by the feminine suffix *-a*, it has nothing to do with feminine biological sex. Native speakers of Arabic usually do not think of singulative nouns as feminine. Therefore, it is relevant to ask if the singulative is an instance of gender. Different views of this matter are present in the literature.

Zabbal (2002), Fassi Fehri (2004),⁹ Borer and Ouwayda (2010) treat the singulative like a marker that is nothing more than *homophonous* to the feminine affix, hence dissociating the two based on their different functions. In such treatments, the singulative is viewed as a morpheme that is completely distinct from gender and is represented on another (higher) functional level. On the other hand, Acquaviva (2008) and Kramer (2015) treat the singulative like other instances of the feminine gender based on the fact that these functions share the same suffix and agreement patterns. They both treat the singulative as a noun-forming affix.¹⁰

To answer the question about whether or not the singulative should be treated as an instance of gender, let us return to our definition of gender, as presented in [Chapter 2](#):

(39) **Definition of gender (standard)**

Genders are classes of nouns reflected in the behavior of associated words.
(Hockett, 1958, p.231)

The key morphological element that helps us recognize gender, according to the definition in (39), is that it is reflected on associated words, through agreement. This element has helped us distinguish gender from noun and numeral classifiers. If we look at the types of gender reviewed in [Chapter 2](#) and the singulative, we can observe that they behave similarly in terms of agreement.

⁹In subsequent work, however, Fassi Fehri treats the singulative as an instance of the feminine but with a different function than the lower gender (Fassi Fehri, 2018a,b)

¹⁰Although, the singulative is merged under a higher *n* in both analyses.

Subsequently, the same affix was used to mark the singulative function, marking the passage from the collective to the individual. This function is, of course, still present and very productive in modern dialects, as profusely discussed, and shown in (42).

(42) tut → tut-a [Tunisian Arabic]
 berry.MASC.COLL berry-FEM.SING
 ‘berries, a berry’

Then, the feminine affix *-a* was used as a marker of group formation in Proto-Semitic and this is what we find in the case of the plurative as discussed in [Section 3.2](#) and in [Chapter 4](#). This group-forming use of the feminine affix is illustrated in (43).

(43) falleḥ → falleḥ-a [Tunisian Arabic]
 farmer.MASC.SG farmer-GR
 ‘a farmer, farmers’

Finally, the affix *-a* became associated with feminine gender and started being used to mark the feminine members of mating pairs (44).

(44) ʕamm ʕamm-a [Tunisian Arabic]
 uncle aunt
 ‘an uncle, and aunt’

Note that all of the uses presented in the examples above not only involve the affix *-a* but also trigger feminine agreement on controlled categories. We are thus

talking about feminine gender in the sense of [Hockett](#)'s definition. In sum, according to the studies by [Hasselbach \(2014a,b\)](#), the historical changes involving the suffix *-a* in Proto-Semitic followed the sequence in (45).

Without getting into the details of how these diachronic changes occurred, what we can retain from this is that the synchronic situation of the feminine marker has traces of all the series of changes of status that this affix went through in time. It is thus unsurprising, considering the historical trajectory of *-a*, to find such variation in roles and functional representations within one and the same morpheme.

3.3 Proposal: interpretable gender on Num

So far, I have discussed the nature of Arabic collective nouns and the role of the singulative which is marked by the feminine affix *-a*. I proposed in [Chapter 2](#), following other authors ([Lecarme, 2002](#); [Kihm, 2005](#); [Lowenstamm, 2008](#); [Acquaviva, 2009](#); [Kramer, 2009, 2014, 2015](#); [Hammerly, 2018](#)) that gender is a nominalizer hosted on *n*. The fact that the singulative operation is marked by gender shift raises questions about the role of gender on the singulative and how it should be represented in the syntax. Is it still hosted on *n*? Is it interpretable? Most analyses of the singulative have looked at the syntactic role of the singulative without conducting a close examination of the facts about gender. In certain cases, the singulative marker is considered to be a *homophone* of gender at most ([Zabbal, 2002](#); [Borer and Ouwayda, 2010](#)), hence avoiding the discussion around the different functions of gender. In such cases, the singulative is said to have a function different from that of gender.

Kramer (2015), in an attempt to regroup all manifestations of gender under the same head (n), treats the singulative as a derivational process and claims that it belongs on a second level on n , where it is treated as uninterpretable gender since it does not translate into biological sex. In the analysis that I adopt, the singulative is viewed as “number in disguise”. *Contra* Acquaviva (2008) and Kramer (2015), I show that the singulative, although marked by gender, is an inflectional operation that must be represented on Num. Gender on the singulative is neither interpretable on n , nor is it arbitrary as is the feminine gender on a French noun such as *la chaise* ‘the chair’. Therefore, it cannot be represented as *uninterpretable* on n . I argued in Chapter 2 that gender can be uninterpretable on Num. I now show that the Arabic singulative is an instance of interpretable gender on Num.

3.3.1 The singulative is inflectional

As already hinted by my observations in the previous sections, I argue against a derivational treatment of the singulative *-a* (as put forward by Acquaviva 2008; Kramer 2015). In this section, I show that the singulative is not a noun-forming process, but rather an inflectional operation. The main arguments for this claim are the high productivity and regularity of the singulative and its clear association with number. Based on these arguments, I conclude that the singulative, although expressed by gender is not hosted on n , since n is a noun-forming head.

In Arabic, the singulative is undoubtedly a productive process, as it can be applied even to neologisms and borrowings, with steady interpretations. Take for example the borrowings in (45), where TA borrowings from French are made singulative by

adding the suffix *-a*.

- (45) a. ananas → ananas-a [Tunisian Arabic]
pineapple.MASC.COLL pineapple-FEM.SING
'pineapples, a pineapple' (borrowing from the French *ananas*)
- b. gazuz → gazuz-a
soft drink.MASC.COLL soft drink-FEM.SING
'soft drink, a (bottle of) soft drink' (borrowing from the French *gazeuse*)

Productivity is characteristic of stable inflectional classes (Wurzel 1984, 1987; Dressler 1998, 117-119; Stolz 2001), and the integration of borrowings into the inflectional system of a given language is a token of productivity. In Acquaviva's (2008, 247) view, the distribution and function of the singulative are, to a great extent, lexical. He claims that the role of the singulative is determined by the semantics of the collective noun, and the set of collective nouns is often too small (Acquaviva, 2015). With regards to this observation, I argue that the collective class is a relatively stable one, in terms of its membership and semantics. First, although the membership of nouns in the collective class is not fixed across dialects in Arabic, the idea of inherent plurality and collective representation is present throughout this whole system, even across languages. We are thus, not talking about random lexical elements taking exceptional singulative marking, but rather a stable class based on solid semantic grounds. Second, the possibility of integration of borrowings and neologisms to the class shows that it is not arbitrary nor too small. On the contrary, this shows that the class is productive and ever growing. Grammars of languages with collective systems are full of examples of collective nouns. Stolz (2001) lists more than seventy

collective nouns for Welsh and this does not appear to be exhaustive. Acquaviva also claims that the meaning of singulatives varies enormously: “it refers to members of collections, atomic parts of granular masses, detached pieces of matter, but also to objects made up of material, or bounded extensions to mass.” The shifting meaning of the singulative is, for [Acquaviva \(2008\)](#), a sign that it involves a derivational process. While there is indeed variation as to the meaning of the singulative, it always involves a part-whole relation with the collective counterpart. The singulative consistently operates on mass nouns and turns them into count, acting like numeral classifiers ([Greenberg, 1972](#); [Zabbal, 2002](#); [Fassi Fehri, 2004, 2012](#); [Mathieu, 2009, 2012, 2014](#); [Ouwayda, 2014](#)). That being said, the singulative is clearly productive and stable in its function and interpretation.

The singulative is associated with number, like any other process turning mass nouns into count. A collective noun cannot be pluralized nor combined with numerals (46-a). Its singulative counterpart, on the other hand, can (46-b).

- (46) a. *xamsa tut. [Tunisian Arabic]
 five berry.MASC.COLL
 ‘five berries’
- b. xamsa tut-a-at.
 five berry-FEM.SING-PL
 ‘five berries’

[Kramer \(2015\)](#) argues that the singulative is an uninterpretable gender feature on *n*. To claim that the singulative is a derivational process under *n* results in the loss of its individuating semantic contribution, which has clear connection with number.

This would amount to treating the gender feature marking the singulative operation like any form of arbitrary gender that has no effect on number interpretation. Moreover, many languages mark the singulative operation with an affix that does not involve gender whatsoever. In Welsh, for example, the singulative operation is not marked by gender shift, but by an affix that varies according to the base gender. In Dagaare, the singulative suffix is an allomorph of the plural marker that takes the singulative interpretation if the base noun is part of the collective class. This strongly suggests that the singulative is more linked to number than to gender and should not receive the same treatment as noun-forming instances of gender.

3.3.2 The singulative: Interpretable gender on Num

I proposed, in [Chapter 2](#), that gender can be, in some instances, expressed on Num (following [Ritter 1993](#)). The idea was that in these cases, gender does not have a nominalizing function, but is rather an inherent feature of the Num head (an idea also pursued by [Farkas 1990](#) for Romanian and [Bernstein 1993](#) for Walloon). Based on the data at hand, I argued that the feminine agreement found on the plural of inanimate nouns in Arabic resulted from an uninterpretable feminine gender feature on Num. In this section, I pursue the idea that the singulative is an expression of interpretable gender on Num.

Based on the arguments presented in the previous sections of this chapter, it appears that the singulative is:

- a) an instance of gender (according to [Hockett's](#) classic definition);
- b) a productive and hence inflectional operation;

c) linked to number.

I also established that whether a noun belongs or not to the collective class, it is not specified on the root, but rather constructed in the grammar of the dialect or language at hand. Although there is a clear connection between inherent plurality and belonging to the collective class, cross-linguistic and dialectal variation indicate that such information originates from a higher level in the syntactic structure. A noun does not come from the lexicon specified as count or collective.

Another important observation is that biological gender distinctions are neutralized in collective nouns, even when the noun refers to an animate entity. Basically, we want to be able to distinguish between the feminine morpheme on *fakruna* ‘a female turtle’ and the one on *jormena* ‘a duck (singulative)’. This can be reached by attributing the right features to the *n* on nouns and more importantly by encoding their interpretability in a specific way.

Based on these observations and the conclusions drawn from them, I first propose that the collective classification of a noun is encoded on the *n* head. This accounts for the fact that such a classification is the result of a morphologization of this category and not of some feature inherent to the lexicon. Moreover, this allows us to put some restrictions on the interpretability of gender on *n*: we want our *n* marked [+COLL] to block any possibility of interpreting gender at this level. Remember that in the count paradigm, the gender of animate nouns is always interpreted on *n*. In the collective paradigm, however, gender is never interpreted on *n*. This prevents us from confusing a singulative noun with a noun referring to a female entity. As a result, we get an additional type of *n* in TA (47-f). The [+COLL] feature neutralizes

the potential semantic contribution of the [-FEM] feature on *n* by signaling that we are in a different paradigm where gender is to be interpreted at a higher level.

(47) **Types of *n* in TA**

- a. *n* *i* [+FEM] Female natural gender
- b. *n* *i* [-FEM] Male natural gender
- c. *n* No natural gender (or natural gender irrelevant/unknown)
- d. *n* *u* [+FEM] Feminine arbitrary gender
- e. *n* *u* [-FEM] Masculine arbitrary gender
- f. *n* [+COLL] Collective base

We therefore obtain the following morpho-semantic mapping:

(48) **Morpho-semantic mapping: ‘turtle’ vs ‘duck’**

- a. [*n*[-FEM] [$\sqrt{\text{fakrun}}$]] = ‘a male turtle’
- b. [*n*[-FEM; +COLL] [$\sqrt{\text{jormen}}$]] = ‘ducks’

(48) accounts for the fact that two nouns denoting animate entities can trigger masculine agreement on targets without referring to the same biological gender, the difference being attributed to their distinct respective morphological structures. In the case of masculine collective nouns, [+COLL] is an expletive gender value and collective nouns thus take the default masculine gender.

Nouns of the collective class are characterized by having their gender features interpreted higher, on Num. The singulative marker *-a* is interpretable on Num,

yielding a unit reading. For collective nouns, the Num head does not project and we therefore get a noun with a mass distribution (as earlier demonstrated). In the absence of any higher gender feature due to the lack of a Num head, the collective feature on *n* is the one triggering masculine agreement on targets. In the case of the singulative, the presence of a higher head marked with feminine gender results in the noun being syntactically feminine (higher gender wins, see [Alexiadou 2004](#); [Steriopolo and Wiltschko 2010](#)).

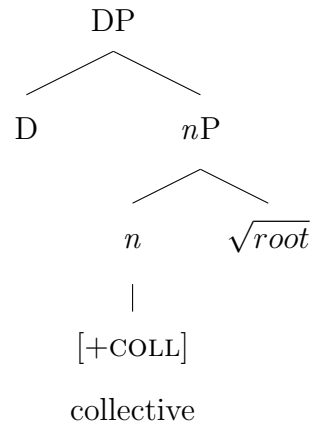
The different flavours of Num now include a new member, as illustrated in (49-d).

(49) **Types of Num in TA**

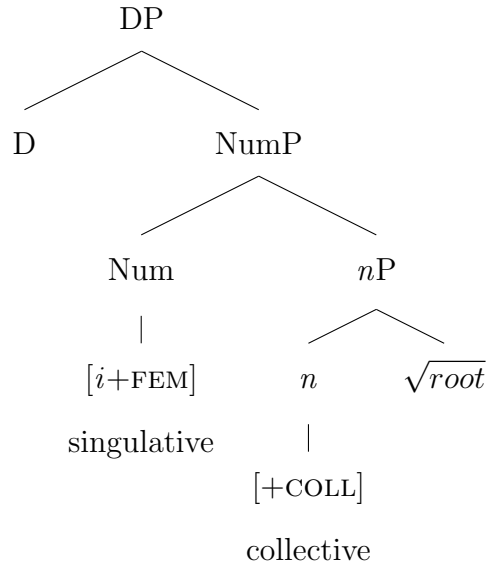
- | | | |
|----|----------------------|-----------------------|
| a. | Num [-PL] | singular nouns |
| b. | Num [+PL] | plural animate nouns |
| c. | Num [+FEM,+PL] | plural inanimate noun |
| d. | Num [<i>i</i> +FEM] | singulative noun |

The resulting syntactic trees for the collective and the singulative are represented in (50-a) and (50-b), respectively.

(50) a.



b.



For languages where the singulative is not marked by gender, the same structure can be assumed, where any affix marking the singulative would be interpreted at the Num level.

In parallel with my previous proposal that gender can be an uninterpretable feature on Num (based on data presented in [Chapter 2](#), where inanimate plurals always take the feminine form), I propose that gender can also be an interpretable feature on

Num. This is the case of the singulative in Arabic and other languages marking the singulative operation by a gender affix. This proposal follows from the hypothesis that the morpheme *-a* has migrated from a derivational to an inflectional operator and vice versa, as put forward by Dali and Mathieu (2019). The present proposal is simply a synchronic account of diachronic facts obtained through reconstruction studies. From this point of view, it seems more natural to analyze the feminine morpheme as occupying different functional levels given its multiple functions.

This analysis has an advantage over accounts where the Num head is used interchangeably with a Class or Gender head. First, it is less confusing as gender and classifier heads can be represented at different levels, namely below *n*P or above it. Second, it eliminates the potential ambiguity between a Classifier head as a marker of division and a Classifier head in the sense of noun classification, which has a function closer to *n*, since it does not instantiate an individualizing function but rather acts as a nominalizer as advanced in [Chapter 2](#). My proposal accounts both for the gender component and the dividing function of the singulative by marking the operation as an interpretable gender feature on Num.

3.3.3 Summary

In this section, I argued that the singulative is an inflectional operation hosted on Num. Although marked by gender, the singulative is not a noun-forming instance of this feature and hence is not on *n*. The gender feature on nouns of the collective class are not interpretable on *n*, since they do not correlate with biological sex. However, it is not arbitrary like gender on inanimate nouns of the count class. The

main arguments in favour of an inflectional-based analysis of the singulative are the productivity of the operation and the consistency of its interpretation, the fact that it is not marked by gender in all languages, its close correlation with number, and the fact that it operates on nouns that are already formed.

I also established that, although closely related to number, the singulative is an instance of gender, based on the classic definition of gender by [Hockett \(1958\)](#), which characterizes it as a way of classifying nouns that is reflected on agreeing elements. I concluded that the singulative is an instance of number in disguise.

I discussed some conclusions drawn from reconstruction studies of Proto-Semitic [Hasselbach \(2014a,b\)](#) that show that the feminine marker *-a* has gone through diachronic changes that are reflected in the synchronic facts. The status of the feminine morpheme went from derivational to inflectional and vice-versa over time ([Dali and Mathieu, 2019b](#)). It is thus unsurprising to find different functions and syntactic levels of gender synchronically, given the diachronic picture presented by the reconstruction studies.

To account for these formal and functional characteristics of the singulative, I proposed an account where gender can be interpretable on Num. In addition to accounting for the facts about the singulative operation, it complements the proposal made in [Chapter 2](#), whereby the Number head can bear uninterpretable gender features. I now turn to previous and alternative analyses alternatives of the singulative, with a focus on the different treatments of the function of gender in singulative systems.

3.4 Previous analyses

The Arabic collective has been the subject of many analyses, both in the descriptive and formal literature. More specifically, different accounts have been offered of the semantic relationship between the collective and the singulative, and the ways to formalize this relationship and represent it in a syntactic structure. In this section, I go over the predominant analyses and comment each one of them in terms of their potential predictions.

3.4.1 Greenberg (1972)

Greenberg's seminal work on language universals includes a substantial contribution to the understanding of classifier systems and their relation with plural marking. His valuable observations of classifier and collective-singulative systems constitute an important foundation for influential linguistic analyses of number marking (Borer, 2005) and the Arabic collective system (Zabbal, 2002; Borer and Owayda, 2010; Fassi Fehri, 2012, 2018a,b; Mathieu, 2012, 2014).

Greenberg's (1972) study was based on a sample of about 100 numeral classifier languages. He advances the following hypotheses:

1. Numeral classifiers involve the overt expression of one kind of quantification, namely, counting by units;
2. The numeral classifier construction is modelled after the measure construction with mass nouns and hence arises in languages with previous mass-count distinction;

3. The classifier in numeral classifier language has the same function as a singulative does in a language with a collective-singulative distinction.

The starting point of his analysis is the observation that numeral classifier languages generally do not have compulsory plural marking on nouns. By way of illustration, consider the case of Mandarin, where the classifier *ben* is required with *shu* ‘book’ with any numeral (51). Note that *book* is in the singular form in both examples in (51).

- (51) a. i ben shu [Mandarin ([Greenberg, 1972](#))]
one CL book.SG
‘one book’
b. san ben shu
three CL book.SG
‘three books’

In non-numeral classifier languages like English, however, plural marking on nouns is obligatory prior to combination with a numeral higher than one (52).

- (52) a. one book
b. three books/*book

The complementary distribution of plural marking and number classifiers suggests that they fill the same function: that of unit counting, as a way of avoiding direct combination with a singular/non-classified noun with numerals. This led [Borer \(2005\)](#) to propose an analysis where plural marking is associated with a functional head, DIV, responsible for “dividing” the noun before it can combine with a numeral.

It appears that languages need some sort of “intermediary” between the numeral and the noun in order for them to combine. Direct combination of these two elements is avoided by means of plural marking or the presence of a numeral classifier. Greenberg notes that in the collective system of Arabic, collective nouns do not combine directly with numerals, but cannot be pluralized to make this combination possible. Either a classifying noun (e.g., *a piece of*, *a portion of*, *a head of*) is interposed and is in direct syntactic construction with the numeral, or the singulative marker is used instead. This suggests that the classifier in a numeral classifier language has the same function as a singulative in a language with a collective-singulative system. To illustrate, let us consider the example in (53) from Omani Arabic.

- (53) a. *thalaath bequd [Omani Arabic (Greenberg, 1972)]
 three gnat.MASC.COLL
 ‘three gnats’
- b. thalaath bequd-aat
 three gnat.FEM.SING.PL
 ‘three gnats’
- c. *thalaath bequd-a
 three gnat-FEM.SING
 ‘three gnats’

Note that the singulative in (53) has to be in its plural form in order for the noun to combine with a numeral. The plural of the singulative is discussed in [Chapter 4](#). Greenberg observes that the Arabic singulative system differs from numeral classifier languages where neither counters nor measurers ever take plural markers, and, unlike typical class systems, the classifiers themselves practically never vary lexically

for number. This fact remains unexplained considering an analysis where numeral classifiers and plural markers are said to fill the same function and hence are expected to be in complementary distribution. This observation sets apart the singulative as a special kind of numeral classifier and has led different authors (Borer and Ouwayda, 2010; Fassi Fehri, 2004, 2012; Mathieu, 2009, 2013) to propose different accounts of the plural and the singulative, as is discussed in [Chapter 4](#).

3.4.2 Ojeda (1992)

Ojeda (1992) adopts a view where Arabic collectives denote kinds, and singulatives denote instances thereof. He assumes that the semantics for Arabic number are based on a set called *the universe of discourse*. Any elements contained in the universe of discourse are said to be *the individuals of the universe of discourse*. It has been argued by Carlson (1977) that both kinds and their instances are individuals of the universe of discourse. This means that any relation between kinds and their instances is a relation between elements of the universe of discourse. Consider in particular *the relation of instantiation*. This is the relation between an individual and a kind just in case the individual is an instance of the kind. For example, the object denoted in *the computer* in (54-a) thus bears the relation of instantiation to the object denoted by *the computer* in (54-b).

- (54) a. Turing repaired the computer.
b. Turing invented the computer.

Turning to the collective nouns of Arabic, the proposal takes the form of the

following constraint on its interpretation:

- (55) Every collective noun denotes, if anything, a singleton subset of the universe of discourse.

Singulative nouns, on the other hand, should denote a set of proper individuals. As lexical items derived from collectives, singulatives denote the set of proper individuals which instantiate the kind denoted by its collective base.

Ojeda also addresses the fact that singulatives can refer to portions rather than to individuals. This is notably the case when the collective base refers to a substance. For instance, the collective *xifab* ‘wood’ in Standard Arabic forms the singulative *xifba* ‘piece of wood’. This type of example questions the primacy of individuation over other partitions. For example, if the singulative can denote a portion reading, it should also be able to denote a portion smaller than a natural unit. However, this is not the case. The singulative always refers to the most salient unit or portion. Ojeda suggests that this primacy of individuation over other partitions may follow from other, extralinguistic considerations. Similar points can be made if we focus on partitions other than individuations, for not all such partitions come equally readily to mind. Some partitions involve “natural kinds” and thus seem more natural than others (e.g., a piece of wood, a loaf of bread). It might thus be thought that singulatives should only be able to denote “natural partitions”. Naturalness, however, is not something that semantics should decide. Natural languages allow us to refer to all kinds of partitions, and make them undenotable is to demand too much of language and too little of other systems of knowledge and belief. Therefore, to an-

swer the question “are partitions other than individuations possible?”, Ojeda would say “yes, singulatives should be able to denote *all* partitions, at least in principle”. Based on this view, Ojeda’s notion of partition is a generalization of the notions of individuation. This point would be conceded if there were instances of singulatives glossed as arbitrary instances, which is unfortunately not the case. For example, we would need examples where the singulative of *beed* ‘egg’ refers to a partition other than an individuation (i.e., an individual egg). To illustrate, let (56-a) be the set of all eggs in the universe of discourse.

- (56) a. {a, b, c}
b. {a+b+c}
c. {b, a+b}
d. {c, a+b}
e. {a, b, c}

(56-b) to (56-e) represent the different partitions of the egg kind, that is, the different ways in which we can divide the instances constituting the egg kind. Assuming Ojeda’s idea that partitions other than individuations are possible, then the singulative should denote, not only (56-e) (the set of individuals), but also any partition from (56-b) to (56-e).

No instance of the singulative based on individualized collectives seems to abound in this sense. If we return to the intuition that collectives can denote both substances and sums of individuals, as demonstrated through the comparative constructions test in the previous section, we can see that if the collective denotes a sum of individuals.

Then the singulative will have no other option than to refer to a natural individual, the most salient partition. For example, it was previously established through the comparative constructions test that *ħut* ‘fish’ is ambiguous between individuated and uninidividuated reference. If Ojeda’s claim is true, the singulative *ħuta* would be able to denote an arbitrary partition of fish. This is not the case (57). If a natural unit is available, then the singulative would denote that unit without fail.

- (57) Kle ħut-a. [Tunisian Arabic]
 ate.he fish-FEM.SING
 ‘He ate a fish.’
- ✓He ate one individual fish.
 x He ate one piece of fish.

To sum up, [Ojeda \(1992\)](#) analyzes Arabic collectives as kinds, and singulatives as refinement functions of instances thereof. Although the singulative always refers to natural units or salient portions, Ojeda claims that nothing prevents it from referring to other arbitrary partitions. The reason why there is no evidence to support this claim is, according to him, based on other, non-linguistic grounds.

3.4.3 [Zabbal \(2002\)](#)

[Zabbal](#)’s view of the collective interpretation diverges from Ojeda’s. Zabbal argues that the collective has a generic as well as a non-generic intepretation. Adopting [Carlson’s \(1977\)](#) semantics of kinds, he proposes that the generic interpretation of the collective is kind-denoting and does not refer to any contextually salient objects.

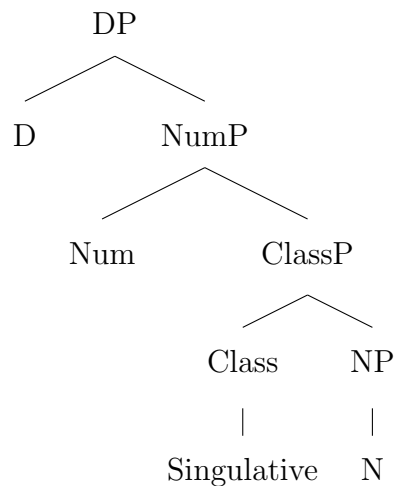
The non-generic interpretation of the collective, on the other hand, has the same semantics as *furniture*-type nouns in English (a semilattice), as demonstrated in the previous section, and refers to particular individuals. This analysis is based on Chierchia's (1998) semantics of plurals and mass nouns, where they are considered to be inherently plural, that is, they come out of the lexicon with plurality already built into their extension. Zabbal unites the generic and non-generic interpretations of the collective by proposing that the collective noun itself is non-generic and that its generic interpretation of the DP comes from the determiner *al* 'the'.

Zabbal's analysis of the function of the singulative suffix *-a* relies on Greenberg's (1972) definition of unit-counters. Singulatives, like unit-counters, combine with a non-count noun (the collective) and the result is a count noun. Zabbal proposes that the semantics of the singulative can be represented as the *singular operator*, SG, which maps a semilattice onto the set of its atoms. The collective has the semantic structure of a semilattice - i.e., it has plurality built in its structure, based on Chierchia's (1998) notion of inherent plurality. A semilattice contains both atoms and sums, therefore it is considered unmarked/neutral for number. Corbett (2000) refers to nouns unmarked for number as having general number. Greenberg (1972) also refers to the "true collectives" of Russian and Arabic as "semantically neither singular nor plural". Semantically, the plural operator, marked by plural morphology, cannot apply to a structure that already contains pluralities as this would result in an empty set. In Chierchia's view, this explains why *furniture*-type nouns cannot have plural morphology. This, however, does not directly predict why numerals cannot combine with *furniture*-type nouns. Chierchia (1998) claims that because the

minimal elements constituting a mass noun are vague, it is not possible to count them. In order to count such nouns, a unit must first be imposed on the structure. Zabbal proposes that numerals cannot compose with structures that contain atoms, only with those that contain only pluralities. However, PL(ural) cannot apply directly to the semilattice, so the atoms must first be isolated from the semilattice. This can be done if we assume an SG operator, which maps the semilattice into a set of its atoms. In English, for mass nouns, this operator is realized as the classifier *piece of*. In Arabic, it is expressed by the singulative suffix.

As for the syntax, Zabbal follows Cheng and Sybesma's (1999) work on the syntax of classifiers in Chinese in assuming a maximal projection *ClassP*, dominated by NumP. He claims that the SG operator is hosted by the head Class, resulting in the structure in (58).

(58)



This analysis relies on the typology of classifiers presented in Greenberg (1972), where the singulative has the role of a classifier in allowing us to count out the denotation

of a noun by units, acting like a unit-counter. The unit counter picks out the units (i.e., atoms) in the denotation of a non-count noun. This is exactly the role of the singulative affix *-a*.

The classifier is base-generated in an intermediate functional head Class of the maximal projection ClassP, which is between NumP and NP. This position is supported by the Arabic facts on number, namely that the singulative may be pluralized (this is discussed in [Chapter 4](#)). In this way, Class must be independent of Num.

Crucially, Zabbal mentions the fact that the singulative *-a* is “homophonous” with the feminine affix, and always triggers feminine agreement. He concludes from this observation that the head Class could be analyzed as gender, which would not be implausible given that both gender and classifiers are a way of classifying nouns. Zabbal also mentions the fact that classifiers are diachronically one of the origins of gender systems ([Greenberg, 1978](#)) without giving further details about the relationship between unit-counters and gender.

Zabbal’s interpretation of the role of the singulative is in line with the intuition that it is needed in order for collective nouns to combine with numerals, hence its function as a counter. However, from my understanding of his proposal, one obstacle of his syntactic treatment of the singulative affix is that it confuses gender/noun class systems with numeral classifiers. We know from the discussion in [Section 2.3.1](#) that as a noun class system, gender is a way of sorting nouns into a small finite set of classes ([Dixon 1982, 1986](#); [Aikhenvald 2000, 19](#)), while numeral classifiers have more to do with allowing the noun to occur in contiguity with numerals. Moreover, according to previous analyses of the function of gender ([Lecarme, 2002](#); [Kihm, 2005](#);

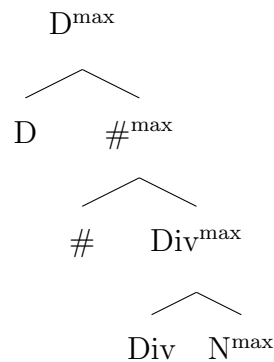
Lowenstamm, 2008; Acquaviva, 2009; Kramer, 2009, 2014, 2015; Hammerly, 2018), gender has a nominalizing function. The singulative, however, plays the role of a unit counter. In the Arabic system, gender undoubtedly marks the singulative function, however, in doing so, it plays a double role, depending on the base to which it is attached. If gender is attached to a count noun, then it functions as a nominalizer, as argued in [Chapter 2](#). However, when attached to a collective noun, gender marks a counting function. In assuming that the Class head is also the head hosting gender in general, Zabbal implies that gender only has one function, which is ambiguous between noun classification and unit counting. Another limitation of this analysis (perhaps because it is not the focus of his study) is that it is not clear from his proposed syntactic structure how the NP is morphologically formed and how one can distinguish between a collective NP and a count NP in the syntax. In his thesis, Zabbal provides structures for the count noun where gender is hosted on the NP. His analysis begs the question: how do we distinguish the role of gender on count nouns from the role of gender on collective nouns?

3.4.4 Borer and Ouwayda (2010)

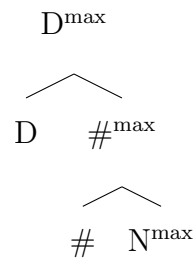
Borer and Ouwayda (2010) treat the Arabic singulative as an instance of DIV, the functional head that also hosts classifiers and plural markers. Their analysis assumes Borer's (2005) exo-skeletal model, where the mass-count distinction is considered purely grammatical rather than lexical. Following Greenberg's (1972) observation that classifiers are in complementary distribution with plural markers (e.g., in classifier languages like Mandarin Chinese), Borer concludes that they both fill

the same function, that of division, and as such, are instantiations of the DIV head. Mass nouns lack the DIVP projection, and this prevents them from combining with numerals. Count and mass structures are presented in (59) and (60).

(59) Count structure:



(60) Mass structure:

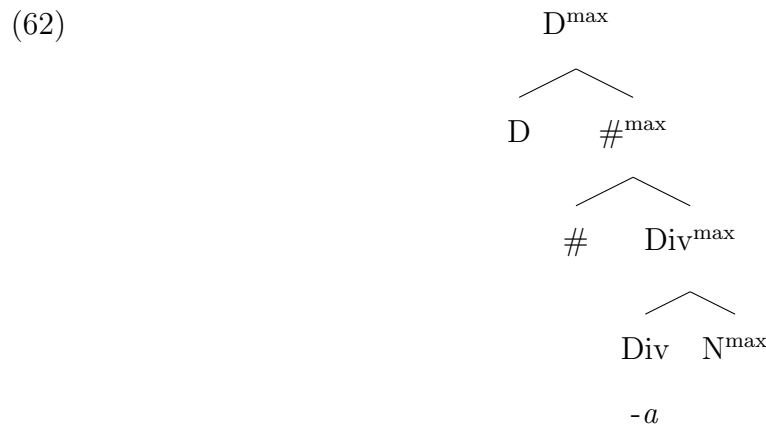


According to this view, the plural interpretation of the plural inflectional marking is nothing but an implicature. Plural marking is not actually plural in its interpretation, but is rather divisional in nature. For instance, [Borer \(2005\)](#) notes that any decimalized number in English must be accompanied by plural inflection (61). Thus, the assumption that the inflection in question is semantically plural becomes questionable.

- (61) 0.2 apples/*apple
 0.1 apples/*apple
 1.0 apples/*apple

When the singulative suffix *-a* is added to a collective nouns, a ‘unit of’ reading emerges. The singulative is then considered to be a classifier ([Greenberg, 1972](#);

Zabbal, 2002; Fassi Fehri, 2004), that is, an instance of DIV. Borer and Ouwayda propose the structure in (62) for the Arabic collective, where the singulative marker is hosted on the DIV head, the head that also hosts classifiers and plural markers.



Having set the theoretical assumptions for their study, Borer and Ouwayda then focus on an apparent counterexample to the prediction that morphological classifiers (such as the singulative marker) and plural marking should be in complementary distribution. In fact, the singulative can in turn be pluralized by the feminine sound plural marker *-at*, and this is problematic given that the singulative and the plural marker are supposed to occupy the same functional head. Borer and Ouwayda circumvent this issue by proposing that *-at* is an instance of agreement with the numeral, since, according to their data (from Lebanese Arabic), the “plural” of the singulative only co-occurs with cardinals. Therefore, it belongs higher in the structure, at the #P level rather than on DIV. This proposal and the plural of the singulative issue are discussed in [Chapter 4](#).

Borer and Ouwayda note that the singulative marker *-a* is homophonous with the

feminine ending *-a* and triggers feminine agreement. However, they claim that the singulative affix is *not* the feminine ending as such, based on the fact that they do not fill the same function. The singulative is a divider, and the feminine is not; when it is compositional, it marks biological feminine gender. It goes without saying that these two exponents with different functions do not belong on the same syntactic head. This is what Borer and Ouwayda propose, by claiming that only the *-a* corresponding to the singulative operation is hosted on DIV.

3.4.5 Fassi Fehri (2004, 2012, 2018a,b)

Fassi Fehri (2004, 2012, 2018a,b) conducted numerous studies about the role of gender and its relation to other functional categories and representations in Arabic. He recognizes the various functions associated with the feminine gender and accounts for these different manifestations by representing the feminine morpheme at different levels in the syntactic structure of the DP. I summarize his proposals in this section.

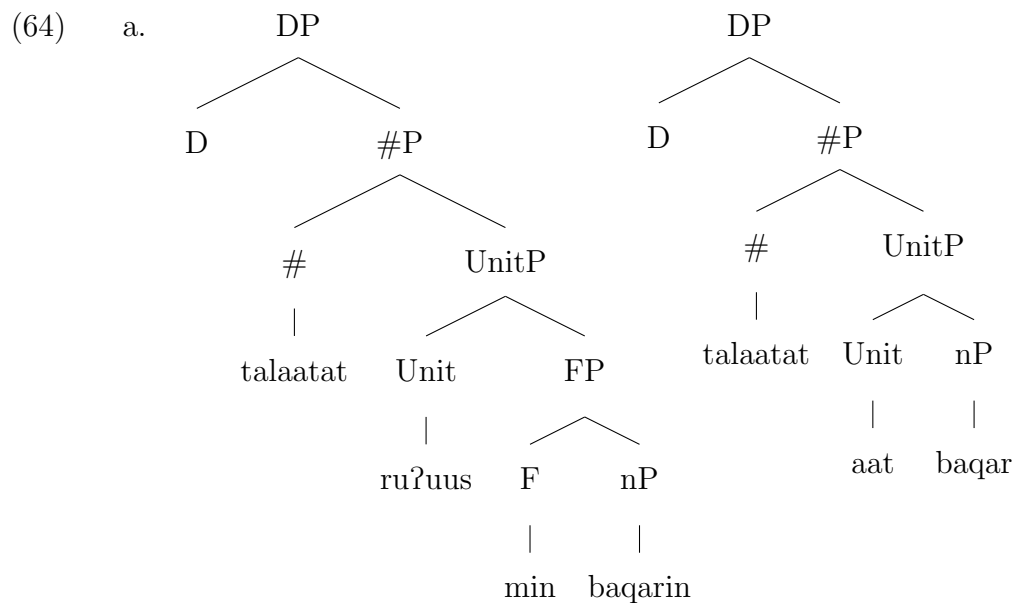
First, note that in earlier works (Fassi Fehri, 2004), Fassi Fehri considered the singulative marker to be a *homophonic* suffix to the feminine gender that functioned as a classifier, forming an “integral unit” out of a collective base. Later on, he attributes multiple roles to the feminine marker, thus treating the singulative not as a homophonous element to the feminine, but as the feminine suffix itself marking a different function (Fassi Fehri, 2018a,b).

Following Greenberg (1972), Zabbal (2002), and Borer and Ouwayda (2010), Fassi Fehri (2018a,b) notes the parallel between the singulative marker and classifiers, based on the fact that both mark the shift from a mass to a count noun. He describes

singulative constructions as an alternative mode of expression of the semi-functional head in indirect noun-headed counting constructions, as illustrated in (63).

- (63) a. talaatat ruʔuus min baqarin [Standard Arabic]
 three head.PL of cattle
 ‘three head of cattle’
- b. tallaatat baqar-a-at
 three cow-FEM.SING-PL
 ‘three cows’

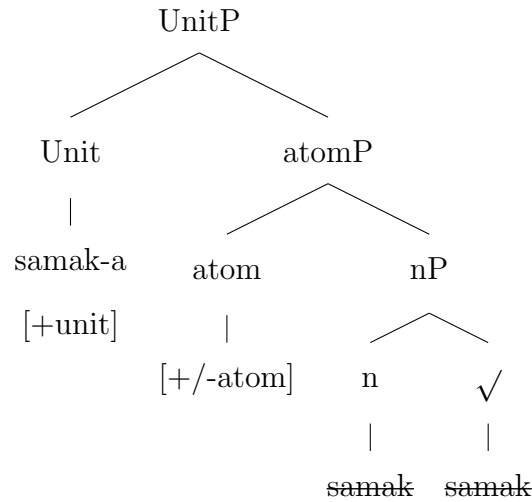
The singulative example in (63-b) is the semantic equivalent of the indirect construction in (63-a). Fassi Fehri (2018a) accounts for this parallel by proposing the constructions in (64), where (64-a) corresponds to (63-a), and (64-b) corresponds to (63-b), the only difference being that the singulative, as a direct unitizer, does not need a Functional Phrase (FP) hosting the lexical PP *min* ‘of’.



The structures in (64) are inspired by [Stickney \(2009\)](#) and [Keenan \(2013\)](#), and UnitP corresponds to [Borer's \(2005\)](#) DivP or Stickney's MP (Measure Phrase). Note that in the tree in (64), it is the plural form of the singulative that is found, *-aat*. This suggests that in this version of his work, Fassi Fehri considers that the plural of the singulative fills the dividing function, thus allowing the noun to combine with a numeral (*talaatat* 'three' in this case).

In a later proposal ([Fassi Fehri, 2018a](#)), Fassi Fehri offers a different structure for the singulative, where the collective noun is dominated by two functional heads with distinct roles: the Atom head divides the collective base into atoms (at this stage, we have a complete semilattice with both atoms and sums thereof), and the Unit head picks an atom out of the complete semi-lattice. In sum, the Atom head acts like a sort of "individualizer" or "divider", then the Unit head act likes a "unitizer". This results in the structure in (65). Note that AtomP has the feature [+/- atom], to account for the fact that it consists of a complete semi-lattice with both individuals and sums, while UnitP is marked [+unit], to account for the fact that it only consists of individuals.

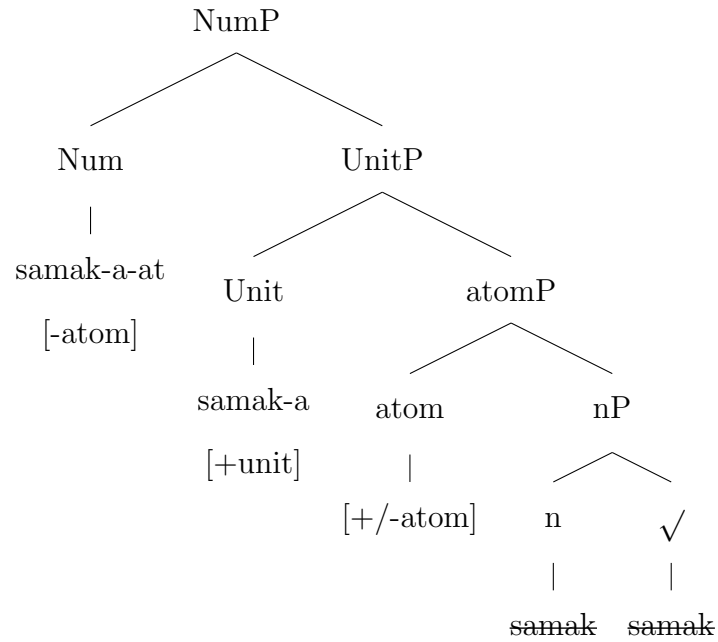
(65)



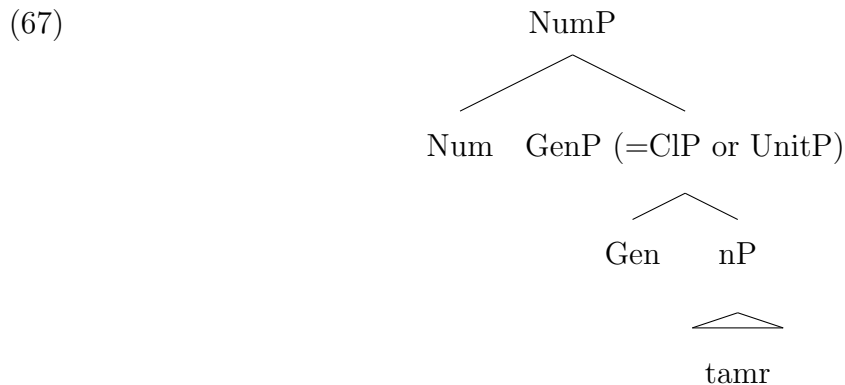
The AtomP projection is, according to Fassi Fehri (2012, 2018a), only present in the structure of singulatives, and not necessary for singulars of count nouns, since their base noun already consists of an individualized semi-lattice. Therefore, Fassi Fehri considers individuation to be encoded within the nP for count nouns, and to be constructed at a higher level for collective nouns.

Contra Borer (2005) and Borer and Ouwayda (2010) but similar to Mathieu (2014), Fassi Fehri (2012; 2018a) argues that there is no complementary distribution between the individualizer (Div, or in this case, Atom) and Num (#), based on the observation that the singulative can co-occur with the plural, resulting in the structure in (66). Note that in (66), the Num head is endowed with a [-atom] feature that removes all the atoms from the complete semi-lattice which consists of individuals and sums, resulting in a lattice consisting exclusively of sums. This analysis was also proposed by (Mathieu, 2014) and is discussed in the next section.

(66)



Finally, in his most recent work, Fassi Fehri (2018a; 2018b) addresses the fact that the feminine can have different functions. He distinguishes the feminine gender with a nominalizing function from the one with an individuating role, emphasizing that the singulative has nothing to do with nominalization. Therefore, he argues that gender as a mark of “nouniness” (on count nouns) is hosted on *n*, while gender with a dividing function is on a higher level, on UnitP. Moreover, he claims that UnitP can be used interchangeably with “ClP” and “GenP” to account for the fact that the singulative is the equivalent of a classifier (in Greenberg’s sense) and is expressed by gender. This analysis is illustrated in the tree in (67).



One potential limitation of such a representation is that while it aims to account for both the gender and the number component of the singulative, it confuses the projections in using the denominations Gen and Cl for the head. If there was to be a Gen head, what would prevent natural or arbitrary gender on count nouns to be hosted on this head, as put forward by [Picallo \(2008\)](#)? Such a representation would predict that a non-singulative language could have gender on two different levels, which should not be the case, given the assumption that gender is supposed to be a nominalizer and hence associated only with the *n* head. In addition, I argued, following [Ritter \(1993\)](#) and [Kramer \(2015\)](#) that the presence of the Gen head is unnecessary and undesirable. Also, how do we account for languages that have a singulative system that does not rely on gender marking? Based on these observations, the gender component of the singulative would best be expressed as a feature under a unified head.

What I retain from Fassi Fehri's research is the multiple treatments and representations of the feminine gender based on the different functions it occupies. Moreover, the splitting of the singulative function into a Atom and a Unit head is an elegant

proposal that efficiently accounts for the individualizing and singulativizing functions of this operation. This proposal is explored in more detail in [Chapter 4](#).

3.4.6 [Acquaviva \(2008\)](#)

[Acquaviva \(2008, 2015\)](#) treats singulatives as the results of a derivational process. The main arguments for this treatment is the fact that singulatives generally involve a gender shift (e.g., in Arabic and Breton), and that they are usually marked by an affix that has multiple functions depending on the lexical base (which is a sign of a lexical operation). Based on data from Breton ([Trépos, 1956](#), 268-9), he claims that the use of the singulative is not limited to collective bases, but can also apply to singular nouns, as in the pair *botez* → *botez-enn* ‘shoe - shoe’, where the second member of the case is said to appear “closer, more material, and more tangible”. This data suggests that the function of the singulative does not only consist in picking discrete entities out of masses, but also into turning abstract objects into indentifiable ones. Based on this function, the singulative is indifferent to the number reference of its nominal base. Moreover, [Acquaviva](#) claims that the singulative does not need a nominal base; it can nominalize adjectives and verbs as in the pair *koant* ‘beautiful’ → *koant-enn* ‘a belle’. The same observation is made for Arabic, where the feminine marker *-a* can turn an adjective into a noun as in *demokrati* ‘democratic’ → *demotkratiy-a* ‘democracy’ (this example is my own). All these observations lead [Acquaviva \(2008, 2015\)](#) to treat the singulative as a noun-forming operation, and not as an inflectional morpheme related to number. In other words, *-a* serves to derive a new noun with a count reading, and not to spell out the singular of the base form.

Acquaviva (2008) acknowledges the contribution of a functional head above the NP as a division operator encoded either through a classifier or grammatical number that enables quantification over the units defined by the division operator (Borer, 2005). However, he suggests that collective nouns, unlike mass, do not lack unity, but rather identity. Therefore, from my understanding of the analysis, the singulative does not partition nouns into units (division), but rather makes the noun reference more “distinguishable”, hence it is not suitable to assign range to the DIV head.

While acknowledging the nominalizing function of the suffix *-a* (and of gender in general, as argued in Chapter 2), I wish to point out that the singulative, although using the same morpheme, has a different function from these uses. Grouping all the instances of the feminine morpheme under the same functional head for the sake of morphemic unification results in overlooking an important generalization, namely, that the singulative is a productive operation with a distinct and regular function. As for the Breton data, it is also possible that the *-enn* affix operating on singular nouns is not an instance of the singulative, but rather a feminine marker with a distinct, derivational function, since number does not seem to be involved. The bottom line is that there is no need to lump all instances of a morpheme under the same head if it has different instances, each filling different functions. Moreover, not all languages with a singulative system use gender distinctions to mark the individualizing function (e.g., Welsh, Dagaare). This weakens the impact of the use of gender as a marker in arguing for a derivational-based analysis of the singulative.

Acquaviva also observes that the singulative has some sort of “packaging function” with verbal nouns, which are not plural in any sense. The result of this deriva-

tion is called *'ism el marra* 'noun of time', *nomen vicis* (literally 'noun of *one* time').

This use is illustrated in (68).

- (68) a. $\text{\textcircled{f}atas}$ 'to sneeze' → $\text{\textcircled{f}at\textcircled{a}s}$ 'sneezing' → $\text{\textcircled{f}atsa}$ 'a sneeze'
b. $\text{\textcircled{b}aas}$ 'to kiss' → $\text{\textcircled{b}oos}$ 'kissing' → $\text{\textcircled{b}oose}$ 'a kiss'
c. $\text{\textcircled{s}aafar}$ 'to travel' → $\text{\textcircled{s}afar}$ 'travelling' → $\text{\textcircled{s}afra}$ 'a trip'
(from [Cowell 1964](#), 302 and [Acquaviva 2008](#), 223)

It is true that the examples in (68) illustrate cases where the singulative operates on verbal nouns. However, I would like to point out that these verbal nouns behave like collective nouns in all aspects. First, they cannot be directly pluralized (69-a) nor combined with numerals (69-b). In addition, they are cumulative (69-c) and can also combine with individuating predicates (69-d).

- (69) a. $\text{\textcircled{f}ats}$ → $*\text{\textcircled{a}\textcircled{f}taas}$ [Tunisian Arabic]
sneezing sneezing.PL
'Sneezing, sneezings'
b. $*\text{\textcircled{x}amsa}$ $\text{\textcircled{f}ats}$
five sneezing
'five sneezings'
c. $\text{\textcircled{b}ar\textcircled{f}a}$ $\text{\textcircled{f}ats}$
lot sneezing
'a lot of sneezing'
d. $\text{\textcircled{h}sebt}$ $\text{\textcircled{el}}$ $\text{\textcircled{f}ats}$.
counted.1SG the sneezing
'I counted the sneezing (sneezes).'

The examples in (69) suggest that deverbal nouns that serve as a base for the singu-

classifying one, as put forth by [Borer \(2005\)](#), and that proper mapping of plurals on the various syntactic categories of the nominal domain is necessary. The idea is that there is a classifying and a counting plural. The classifying plural has an inclusive interpretation in the relevant contexts, while the counting plural has an exclusive one. The folk view and the traditional perspective of the plural is that it refers to *more than one* entities (the exclusive reading), as opposed to the singular. However, a more recent trend in the semantic literature ([Krifka, 1989](#); [Sauerland, 2003](#); [Sauerland et al., 2005](#); [Spector, 2007](#); [Zweig, 2009](#); [Bale et al., 2011](#)) has analyzed the plural as referring to *one or more* (the inclusive reading, in the sense that it includes reference to the singular).

Contra [Borer and Ouwayda \(2010\)](#), [Mathieu \(2013\)](#) argues that while the singular is in complementary distribution with the (inclusive) plural with a classifying function, it can combine with the (exclusive) plural with a counting function (hosted on a higher level). In sum, Mathieu proposes a non-unified analysis of the plural and accounts for the different functions of the plural by distributing it along the nominal spine.

It has been argued that the plural in English has an inclusive interpretation in questions and downward-entailing environments ([Krifka, 1989](#)). For example, the question in (72) can be answered with *Yes, I have one child* or *Yes, I have two children* despite the use of a plural nominal in the question.

(72) Do you have *children*?

[Mathieu \(2013\)](#) argues that Arabic has both inclusive and exclusive plurals. To

illustrate, let us examine the interpretation of the plural of count nouns (73) versus the interpretation of the plural of the singulative (74) in questions and downward-entailing contexts. Note that only the plural with an inclusive reading is expected to fit in such environments, since they have general reference, as opposed to the singular, which refers strictly to individuals. The interrogative in (73) can be answered by *Yes, I met three* or *Yes, I met one*. (74) involves a negation (a downward-entailing context) and the proposition is false if I met two teachers or more but also if I met only one teacher. Therefore, the plural of count nouns has an inclusive interpretation in questions and downward-entailing contexts, as in English.

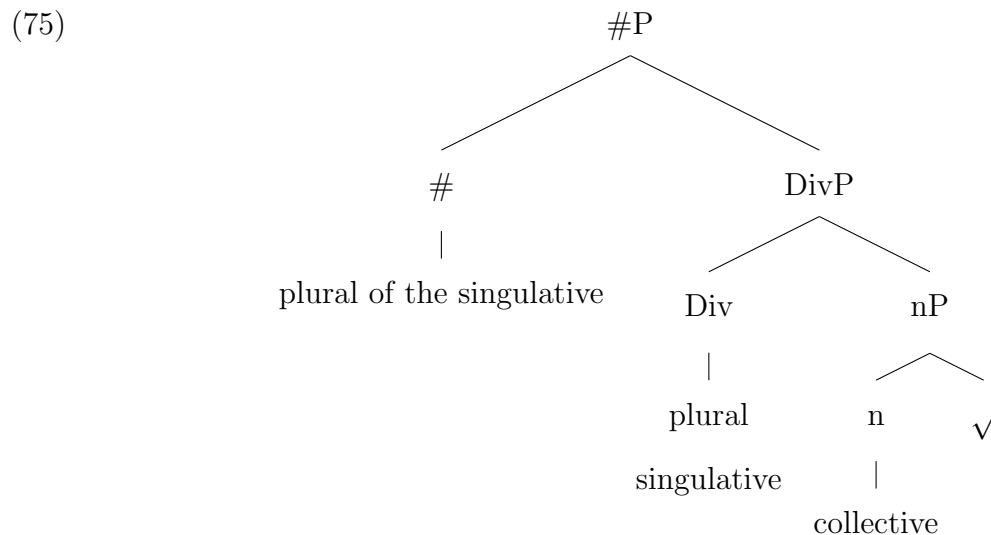
- (73) a. Hal qaabalt mudarris-een? [Saudi Arabic (Mathieu 2013)]
 Q met.you teacher-MASC.PL
 ‘Did you meet teachers?’
- b. Maa qaabaltu mudarris-een.
 not met.I teacher-MASC.PL
 ‘I did not meet teachers.’

However, in the same contexts, the plural of the singulative has a different reading. Consider the contexts in (74). The noun *burtogaal* ‘orange’ is a collective base from Saudi Arabic. It has a singulative counterpart, *burtogaala*, that can be pluralized to yield *burtogaalaat*. While the collective form is inclusive and has more general reference, the plural of the singulative is not neutral enough to be in an interrogative: the noun has been individualized and then pluralized, thus no longer refers to the general concept of oranges. The same pattern is observed in negative contexts, where

the use of the collective form is grammatical and the sentence with a plural of the singulative is ill-formed. This suggests that unlike the collective, the plural of the singulative has an exclusive reading in the relevant contexts.

- (74) a. Hal ʕindik /burtogaal/*burtogaal-a-at? [Saudi Arabic]
 Q have.you orange.MASC.COLL/orange-FEM.SG-PL
 ‘Do you have oranges?’
- b. Ana ma ʔakalt burtogaal/*burtogaal-a-at.
 I not ate.I orange.MASC.COLL/orange-FEM.SG-PL
 ‘I did not eat oranges.’

To account for these different readings, [Mathieu \(2013\)](#) proposes a structure where the inclusive, classifying plural is on the DIV head and the exclusive, counting plural is higher, on #, as illustrated in (75). Therefore, nothing prevents the higher counting plural to appear in conjunction with the singulative, since they do not compete for the same syntactic position.



His account is thus in line with [Zabbal \(2002\)](#) and [Fassi Fehri \(2004, 2012, 2018a\)](#) more so than accounts such as [Acquaviva \(2008\)](#) and [Kramer \(2015\)](#), among others, where the singulative is derivational and is realized lower in the structure (under *n*).

3.4.8 [Kramer \(2015\)](#)

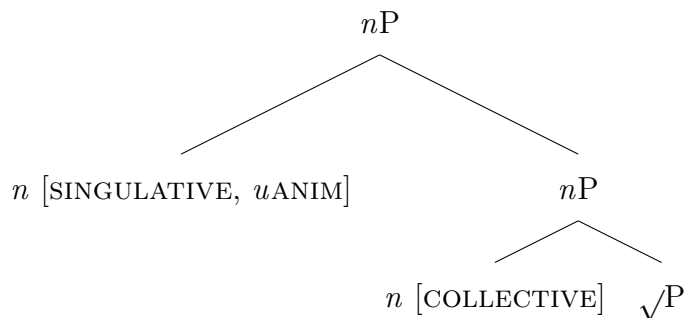
In her influential work on gender, [Kramer \(2009; 2014; 2015\)](#) argues for a unified analysis of gender, where all of its instances are under a single locus, namely *n*. In sum, all instances of gender are said to have a nominalizing function, whether they are interpretable in some sense or not.

To illustrate her claims, [Kramer](#) uses data from [Fox \(Goddard, 2002, 23\)](#) and [Ojibwe \(Mathieu, 2012\)](#), where all collective nouns are inanimate and all singulative nouns are marked with the animate gender, as shown in (76).

- (76) a. zhooniyaah-i ‘silver, money-INANIM’ zhooniyaah-a ‘a coin, a bill-ANIM’
b. owiiyaas-i ‘meat, flesh-INANIM’ a piece of meat-ANIM’
c. anakehkw-i ‘bark-INANIM’ anakehkw-a ‘a piece of bark-ANIM’
(Fox, [Goddard 2002](#), 213)

In the examples in (76), all singulative nouns are marked [ANIM] but refer to inanimate entities. Consequently, [Kramer](#) argues that singulatives involving a shift in gender have uninterpretable gender on *n*. She proposes that the singulative is derived by adding a *n* with semantic features corresponding to a singulative/mass *n*P, as shown in (77).

(77)



Although Kramer proposes that singulatives are generally considered to be part of a language's number system (i.e., they are included in Corbett's (2000) typological survey of number), she argues that they are a flavour of *n*. In fact, she draws a parallel between singulative nouns and Somali plurals. Remember that, as discussed in Chapter 2, Somali plurals involve gender polarity and are thus considered to be hosted on *n* by Lecarme (2002) and Kramer (2015). Kramer notes that Arabic singulatives do not attach to all collective nouns in Classical Arabic (1972, 20). These selectional restrictions of the singulative, in addition to the fact that gender marking is involved with number, are reminiscent of the Somali plurals. Therefore, Kramer submits that singulatives have the same structure: a *n* attaching to a *nP*.

It is unclear how the fact that the distribution of collective nouns shifts from mass to count when they become singulative is accounted for in this analysis. If a derivational operation is responsible for such a change in the countability properties of a noun, then which head encodes this shift, if not a head related to number? This, I sustain, is an important limitation of an *n*-based analysis of the singulative. As for Classical Arabic nouns without a singulative (Greenberg refers to the collective noun *ʔibl* 'camels'), it may be a singular noun referring to a collection of entities.

However, if a noun does not derive a singulative, it means that it is not part of the collective morphological class, assuming that the singulative is the hallmark of the morphological class. First, such examples are extremely rare (Greenberg, in fact, only mentions this one). Also, I showed previously that there are object-mass nouns in Arabic, which have a distribution similar to that of collectives, but which are nonetheless not part of this class. I sustain that the singulative can attach to all collective nouns and beyond, since they are highly productive and can even apply to neologisms and borrowings. Finally, even though singulative is marked by gender shift, this is not the case in all singulative languages (e.g., Dagaare, Welsh). I showed that there is an alternative way to represent the operation while accounting for all of these characteristics and for languages whose singulative system does not involve gender.

3.5 Conclusion

This chapter provided a complete overview of the collective system in TA and across languages, with a focus on its relation to gender marking. Based on some of their distributive and semantic characteristics, I established that collective nouns form a coherent morphological class that conditions the way gender features are interpreted and their locus of representation. The most crucial point made in this chapter is that the singulative, as an inflectional operation marked by the feminine, is marked by interpretable gender on Num. This view stands out from previous *n*-based analyses of the singulative (Acquaviva, 2008; Kramer, 2015). The proposal

put forth here has some theoretical implications that are discussed in more detail in the remainder of the thesis, namely, the fact that one and the same morpheme can have different interpretations depending on the basis to which they attach. This phenomenon, I argue, is ubiquitous in Arabic and is rooted in the diachronic facts about this language and the coexistence of two number systems.

Chapter 4

Tunisian Arabic plurals

4.1 Introduction

This chapter examines the different plural shapes of TA. More specifically, I discuss the different analyses of the broken plural proposed in the literature, as well as their associated syntactic projections. I show that in regular contexts, the broken plural has the same denotation as the sound plural in TA Section 4.2. Therefore, I argue that both plural shapes are inflectional. I also discuss two phenomena related to the topic of plural marking interpretation in TA.

The first phenomenon [Section 4.3](#) explored stems from the observation that some nouns can take both plural shapes in TA. This is unexpected since plural shapes are traditionally said to be attributed to nouns based on phonological information. I show that this alternation is motivated by semantic factors. The main claim is that when the sound plural is used on canonical stems (the phonological context of

broken plurals), the resulting plural denotes ‘a few entities’. Consequently, I argue that the suffix *-at* on canonical nouns is a paucal marker, not a plural. Considering the similarities between this paucal and the one obtained when *-at* is affixed to a singulative, I propose that the use of a paucal marker on nouns of the count class result from a reclassification of a count noun into the collective category.

The second phenomenon (Section 4.4) is related to the observation that broken plurals can trigger feminine singular agreement on targets. Different analyses have discussed various angles of this peculiar agreement pattern (Belnap, 1991; Brustad, 2000; Zabbal, 2002; Dali, 2015; Dali and Mathieu, 2016, 2020; Kramer and Winchester, 2018). What stands out is that this pattern is associated with group interpretation. I argue, following Dali (2015) and Dali and Mathieu (2020), that in these instances, the broken plural is a hybrid noun. It is underlyingly feminine, with the feminine feature expressing group formation as one of its multiple semantic realizations.

4.2 Arabic plurals

So far, I have presented some examples with sound plurals, and others with broken plurals, without explicitly discussing their formal differences and their respective contexts. This section describes the different plural shapes of Arabic and defines both their phonological, syntactic and semantic contexts of use.

4.2.1 Sound and broken plurals

Arabic has two plural shapes: the broken and the sound plural. The broken plural is a pluralization pattern, which consists of reshaping the singular stem, as in (1). This pattern occurs in both nouns (1) and adjectives (2).

- (1) kalb → kleb [Tunisian Arabic]
dog.SG dog.PL
'dog dogs'
- (2) mrið → morða
sick.MASC.PL sick.PL
'sick (sg.), sick (pl.)'

The term “broken” (in Arabic *mukassar*) invokes the idea of affecting the formal integrity of the stem. An alternative pluralization strategy consists of adding a suffix to the singular stem, which remains unchanged, or “sound”. This is what we call “the sound plural”. The suffix *-een* marks the plural of masculine nouns (3-a), and the suffix *-at* (or *-et*, in certain dialects) marks the plural of feminine nouns (3-b).

- (3) a. saḥafi → saḥafiy-*een* [Tunisian Arabic]
journalist.MASC.SG journalist-MASC.PL
'A male journalist, male journalists'
- b. saḥafiy-a → saḥafiya-*at*
journalist-FEM.SG journalist-FEM.PL
'A female journalist, female journalists'

The existence of two formal plural variants in Arabic raises many questions,

namely about the nature of the criteria conditioning the realization of each shape, and whether the different plural shapes should receive a unified or a distinct syntactic treatment. Over the years, a considerable amount of research has been devoted to these questions. In what follows, I discuss the outcomes of these analyses.

First, let us address the question of the contexts conditioning the choice of a plural shape over another. There is a general consensus in the literature that the basis of selection of a plural shape is of a phonological nature. The broken plural, as a morphological process that involves stem internal modification, requires the stem to have a specific shape or pattern in order to modify it. [McCarthy and Prince \(1990a\)](#) have defined this specific shape by a Minimal and a Maximal Stem Constraint. A canonical Arabic stem has a minimum of two moras and a maximum of two syllables. Any stem that goes beyond these limits is considered noncanonical. Noncanonical nouns usually originate from borrowings and are never created by the root-based templatic morphology that we typically find in Semitic languages. [McCarthy and Prince \(1990b\)](#) claim that noncanonical nouns do not normally contribute their root to further derivational processes (e.g., denominal verbs are almost never created from noncanonical forms - see also [Arad 2005](#), ch.2, for a discussion about the internal modification of syllabic roots in Hebrew verbs that originated from borrowings). Furthermore, with rare exceptions, noncanonical loans do not participate in broken plural formation. It is the noncanonicity of these words that prevents them from forming broken plurals, not their status as loans. Many loans are canonically-shaped and can form broken plurals (e.g., *film-aflaam* ‘movie’; *bank-bunuwuk* ‘bank’). In sum, any stem that goes beyond the Minimal and Maximal Stem Constraints is considered

noncanonical and as such, cannot be subject to internal pluralization. Instead, it undergoes the suffixational strategy (i.e., sound plural). [Ratliffe \(1998\)](#) also observed that in all Semitic languages (where these two methods of pluralization are available), the external plural is obligatory for derived nouns. The sound plural is therefore systematically found with transparently derived nouns or adjectives such as participles, deverbals, and diminutives, in addition to noncanonical and unassimilated loans.

Based on the observations about the criteria of stem canonicity, it seems that broken plurals are the default strategy and that few exceptional items (i.e., derived nouns and uncanonical nouns) escape this process to end up with the sound plural suffix. [McCarthy and Prince \(1990b\)](#) conclude that for the Arabic lexicon as a whole, broken plural formation is, by far, the norm rather than the exception. [Boudelaa and Gaskell \(2002\)](#) refute this claim by arguing that the Arabic sound plural is the default based on both qualitative and quantitative criteria. On the qualitative side, they claim that both broken and plurals are productive in the sense that they apply to a singular particular form only if it satisfies certain constraints. For example, broken pluralization requires a certain stem shape, as discussed above. Masculine sound pluralization requires the singular input to be masculine and animate (see [Chapter 2](#) for a discussion about animate nouns taking a feminine sound plural suffix in TA). On the quantitative side, Boudelaa and Gaskell supported their claim by analyzing all nouns listed in the “Basic Lexicon of Modern Standard Arabic” (BLMSA), which consists of the 3000 most frequent words in the language. The authors have extracted a total of 1670 nominal forms out of the corpus, out of which 59% take a sound plural and 41% take a broken plural. They note that for almost every form that has

a broken plural, there is also a corresponding sound plural shape, but the reverse is not true. For example, the noun *qird* ‘monkey’ has a broken plural *quruud*, but its feminine form *qirda* ‘female monkey’ has a sound plural form *qirdaat*.¹ This observation would indicate that the sound plural is quantitatively more important than mere exceptions, as suggested by [McCarthy and Prince \(1990b\)](#) and [Ratliffe \(1998\)](#). Therefore, [Boudelaa and Gaskell \(2002\)](#) conclude that the sound plural is default in Arabic.

The second question relates to the syntactic treatment of each plural shape. Given the stem transformational nature of the broken pluralization process and the fact that this process occurs in the context of root and pattern systems, the following question follows: does the formation of broken plurals result from root and template morphology mappings or are they derived from the singular? The answer to this specific question can further indicate whether the broken plural should be considered as a derivational process (on *n*) like any other category-forming morphological process, or alternatively, as an inflectional operation (on Num) as expected of a form affecting the semantic denotation of a noun in terms of number. Again, the views in the literature are divided about this matter. I explore the different arguments supporting each view in what follows.

The predominant view about the broken plural, and about plural formation in general, from a crosslinguistic perspective, is that the plural is both morphologically

¹Note that in the case of the plural of *qird* ‘monkey’, only the feminine sound plural is possible, and not the masculine sound plural. For instance, *qirdeen* ‘male monkeys’ is not possible. One possible explanation is that the sound plural form is used here as a strategy to specify that we are referring to several female monkeys (the marked gender). This cannot be done through the broken plural shape, which is unmarked for gender.

and semantically constructed on a singular base. Several studies of the broken plural established the form of the singular as the principal factor determining the form of the broken plural in Arabic (Murtonen, 1964; Levy, 1971). Hammond (1988) argues that melodic transfer is also appropriate for the description of root and pattern morphology. He shows, in particular, that a broken plural is better analyzed in terms of melodic transfer from the singular shape than by insertion of a vocalic pattern directly to the root. His claims are based on the following arguments. First, the length of the final vowel of a broken plural shape depends on vowel length in the singular. Second, the spreading of consonants in the plural is based on consonantal spreading in the singular. Basically, if a consonant is spread in the singular, it is also spread in the plural. A third problem with treating broken plural formation in terms of a standard template is the insertion of [w] in the broken plural shape when there is a long vowel in the first syllable of the singular. In sum, vowel length and consonantal spreading are not qualities of the root and as such, should not be visible to mechanisms that directly associate melodic material to the plural template. Hammond attributes these three phenomena to melodic transfer rather than vocalic insertion to a consonantal root pattern. Based on this approach, the issues listed above are accounted for, since transfer allows plural formation to see the singular template where vowel length and consonantal spreading are represented.

In the same vein, McCarthy and Prince (1990b) claim that all the properties of a canonical stem are carried over from singular to plural, despite the fact that the root itself contains no information about canonical form. McCarthy and Prince (1990a) have independently defined the canonical stem for Arabic, based on the Prosodic

Morphology Hypothesis. As mentioned above, the canonical stem has a minimum of two moras and a maximum of two syllables. These constraints include notions like vowel quantity, number of syllables and consonant spreading. The latter notions define exactly the kind of information that the root abstracts away from. In a true templatic morphology setting, only the root consonantism is carried over from one form to another in a prosodically diverse set (e.g., *kitaab* ‘book’, *kaateb* ‘writer’, *kataba* ‘he wrote’). The broken plural, then, cannot be obtained with the ordinary resources of root and template morphology. [McCarthy and Prince \(1990b\)](#) also add more supporting evidence for a melodic transfer treatment of the broken plural (from the singular) based on two arguments. First, they observe that some derivational affixes like *m* in *miftaaḥ* ‘key’, derived from the consonantal root /f t ḥ/ that carries the concept of ‘opening’ are transferred intact. This indicates that the root is also morphologically inappropriate as the basis of broken plural formation. Moreover, they discuss the case of the “plural of the plural”, in which a broken plural is formed from a stem that is itself a broken plural. In the second plural form, the final syllable vowel length is transferred from the (already plural) base. This vowel length comes from the prior pluralization process, not from the singular, and much less from the consonantal root.

[Kihm \(2003\)](#) presents a diverging opinion from the one attributing broken plural formation to melodic transfer from a singular base. In his view, broken plurals are not derived from the corresponding singulars but both forms represent distinct realizations, although related, since they emerge from the same root. He enumerates a list of characteristics highlighting the derivational nature of broken plurals, more

specifically, the multiplicity and imprevisibility of these forms. It appears that one and the same noun can have more than one corresponding broken plural shape. Moreover, he claims that one cannot necessarily predict the shape of a broken plural based on the singular. For instance, he argues that iambicity is not univocally associated to plurality (e.g., *kilaab* ‘dogs’ is plural, while the prosodically similar *kitaab* ‘books’ is singular). This suggests that broken plural patterns are very ‘choosy’ of their roots, which is very characteristic of derivational processes (see [Lecarme 2002](#) for similar claims about the plural in Somali).

Naturally, whether internal pluralization is a process applying to the root or to a singular stem raises the question of derivational versus inflectional operation. Assuming that the suffixational plural is inflectional in nature, since we are talking about an affix dedicated to number interpretation, the question is whether the broken plural should receive the same treatment. From a morphological point of view, [Kihm \(2003\)](#) clearly states that broken plurals are derivational in nature. They simply categorize the root, just like nominalization, and should therefore be hosted on *n*. [Lahrouchi and Lampitelli \(2014\)](#) adopt a similar view for Moroccan Arabic, advancing both morphological and semantic arguments. On the morphological side, they provide arguments similar to those presented by [Lecarme \(2002\)](#) for Somali, and [Kihm \(2003\)](#), namely, the fact that unlike sound plurals, broken plural formation is more restrictive in terms of the roots it applies to. They also mention the fact that double pluralization² is always applied on a broken plural base, which means that the latter should be closer to the root than the sound plural is. Moreover, they

²Double pluralization is very rare in TA, but they also look at other languages like Amharic, where it is more productive.

add that there is a crucial semantic difference between sound and broken plurals in Moroccan Arabic. The difference resides in the observation that the use of a sound plural indicates a definite number, usually occurring with numerals, whereas the corresponding broken plural has a collective reading. The following example (4) is provided as supporting evidence.

- (4) a. mmul l-kwast [Moroccan Arabic]
owner DEF-tape.FEM.PL
'tape seller, record store'
- b. zuʒ kaset-at
two tape.FEM-PL
'two tapes'

Lahrouchi and Lampitelli's (2014) claim is that the broken plural is used as a collective, in more general contexts like the one in (4-a), while the sound plural is used in contexts with a definite number like (4-b) (see also Lahrouchi and Ridouane 2016 for similar arguments). Based on the observation that broken and sound plurals of Arabic are different in both morphological and semantic perspectives, they argue that these plurals are not hosted on the same syntactic head; the broken plural is on *n*, and the sound plural, as a true number exponent, is on Num.³

Finally, Acquaviva (2008) also addresses the question of whether the Arabic broken plural is a lexical/derivational or an inflectional one like the sound plural. By doing so, he raises a point about the importance of distinguishing formal and semantic properties when evaluating the status of a plural shape. In his view, Arabic

³In Section 4.4, I address the semantic differences between broken sound plurals. I argue that these differences are not systematic but rather they occur when the plural forms are used contrastively.

broken plurals are not morphemes that spell out plural number. This suggests that they are lexical. However, putting their form aside and considering only their semantic characteristics, broken plurals fit entirely within the inflectional number system since they have no special semantic meaning (their denotation and context of use are the same as those of sound plurals). Acquaviva's claims are based on the following arguments.

Let us discuss the morphological aspects. The relationship between singulars and broken plurals is not one-to-one, but rather one-to-many, suggesting that internal pluralization results from the interaction of both grammatical and non-grammatical knowledge. This involves some degree of arbitrariness and learning, which suggests that internal pluralization is part of a lexical process rather than an inflectional one. Moreover, Acquaviva discusses that some patterns used for the plural of some nouns appear in the singular of other nouns (this was also later reported by [Kihm 2003](#), as discussed above, with the prosodically similar pair *kilaab* 'dogs' - *kitaab* 'book'). Finally, a stronger piece of evidence that broken plural shapes do not signify plurality *per se*, is that they can themselves feed pluralization. For example, the noun *balad* 'village' takes the plural shape *bilaad* that is translated as 'country'. The latter form is syntactically singular and can feed broken plural formation, to give *buldaan* 'countries'. According to Acquaviva, based on its properties, the broken plural appears to be the output of a stem-forming process, that he also refers to as "Level I operation" (see also [Ratcliffe 1998](#) and [Kihm 2003](#) for similar conclusions).

I now turn to the semantic aspect of broken plurals. Acquaviva first observes that broken plurals have no special meaning and are unambiguously the equivalents

of sound plurals when it comes to their distribution. Since broken plurals have no semantic distinctiveness as a morphological class, they cannot be considered lexical plurals in the sense of having a morphology that expresses a special reading. Another argument against the lexical treatment of broken plurals put forth by Acquaviva is the fact that they can apply to adjectives as much as to nouns. Since adjectives are agreement targets, not controllers, this observations indicates that the broken plural operation is an inflectional one. Considering how broken plural patterns apply to adjectives, a stem-forming process analysis of this plural shape would be very unlikely. Adjectives do not typically get number information as part of their derivational formation (which is attributed to the inflectional head ‘a(djective)’ in the DM framework), but rather acquire this inflectional information through agreement.

The third argument against a lexical analysis of the broken plural is based on the previous claim that broken plurals can feed second level broken plural formation. I discussed the case of the broken plural *bilaad* ‘cities’, that can be reanalyzed as a singular meaning ‘country’ and further repluralized as *buldaan* ‘countries’. At first glance, it seems like broken plural formation can lexicalize a plural form out of its original inflectional paradigm. But we can also see that this is part of a larger pattern whereby the input for broken pluralization comes from the reanalysis of the output of a stem-changing operation. This suggests that the reanalysis of ‘cities’ as ‘country’ is not determined by broken plural formation, but can generally happen to forms generated by other stem-changing (or Level I) operations. It is not broken plural formation that leads to lexicalization, understood as the reanalysis of a grammatical form of a word as a related but distinct word. What happens instead is that a

Level I-derived stem may get lexicalized, and this can involve broken plurals. Based on these arguments, Acquaviva concludes that broken plurals are the products of stem-forming processes, but that they still serve as exponents of a perfectly regular category of inflectional number.

In this section, I have presented different views concerning the broken plural, namely about its phonological environment, its morphological nature and its semantic contribution. I now turn to my own analysis of the broken plural, which, as I argue, is syntactically similar to the sound plural in regular contexts. I argue that the broken plural pattern is an inflectional one since it is associated to information about number. My claim is further supported by novel data in [Section 4.3](#), which shows that sound and broken plurals only differ in special contexts.

4.2.2 Broken and sound plurals are equivalent

There are two opposing perspectives of the plural in the semantic literature. The *Strong Theory* ([Greenberg, 1966](#); [De Swart and Zwart, 2009](#)) is in line with the traditional view that the plural is both morphologically and semantically marked with respect to the singular. In this approach, the plural is considered to be exclusive, referring only to sums, and not to individuals. In the *Weak Theory*, ([Krifka, 1989](#); [Sauerland, 2003](#); [Sauerland et al., 2005](#); [Spector, 2007](#); [Zweig, 2009](#); [Bale et al., 2011](#)), the plural includes reference to individuals in the relevant contexts. In this view, the plural has more general reference, therefore it is semantically *unmarked*. In this section, I test the contexts of use for both plural shapes in TA and show that they are semantically equivalent.

Proponents of the *Strong Theory* of the plural (Greenberg, 1966; De Swart and Zwart, 2009) associate plural marking exclusively to sum values. This view is in line with Horn's (1984) Division of Pragmatic Labor, where marked forms go with marked meanings. In De Swart and Zwart's (2009) account, syntactic projections are only present when associated with morphological materials. Hence, number inflection is necessary for a nominal to have a Num projection. This view is supported by data from Hungarian incorporated singulars (5) and Dutch bare predicate nominals (6), both having singular or plural reference.

- (5) Mari bélyeget gyüt. [Hungarian]
 Mari stamp.ACC collect
 'Mary is collecting stamps.'
- (6) Jan en Sofie zijn leraar. [Dutch]
 Jan and Sophie are teacher
 'Jan and Sophie are teachers.'

De Swart and Zwart attribute this inclusiveness of reference to the lack of number marking, leading to the absence of a numeral projection. A special meaning only arises if it is morphologically marked. Therefore, only plural marking can trigger the projection of a Num head, and, by extension, give rise to special semantics. Along the same lines, Greenberg (1966) claims that plural categories are always semantically and morphologically marked in relation to the singular, as there is no language where the plural is not marked by some overt allomorph.

In the Weak Theory (Krifka, 1989; Sauerland, 2003; Sauerland et al., 2005; Spector, 2007; Zweig, 2009; Bale et al., 2011), the plural is considered to have no inherent

number presupposition, and hence, to have a weaker meaning than the singular. In this view, the plural is semantically unmarked, as it has mixed reference, including both the plural and the singular meanings. The correlation between morphological and semantic markedness is negative (Bale et al., 2011), since the plural is morphologically marked, with a specific morpheme associated to plural meaning (e.g., *-s* in English). For instance, Sauerland et al. (2005) show that indefinite plurals do not exclude singularity in questions and downward entailing environments. Consider the examples in (7), where plural marking occurs in the scope of negation (7-a) and in a question (7-b).

- (7) a. Kai has found no eggs.
b. Does your office have windows?

The sentence in (7-a) is judged false if Kai has found a single egg, and (7-b) can be answered with ‘yes, one.’ This shows that indefinite plurals in questions and downward entailing contexts include reference to individuals. Along these lines, Krifka (1989) shows that bare plurals can refer to one object or even less, as in (8).

- (8) Did you eat apples?
Yes, I ate half an apple./#No, I ate (only) half an apple.

In (8), *apples* refers to less than one entity. Consequently, Krifka claims that the only restriction applying to the use of bare plurals in English is that it must refer to ‘more than zero’, and not to ‘more than one’, as is generally assumed and claimed by the proponents of the Stong Theory. Following this observation, he questions the

use of the singular form in cases where there is evidence that the number of entities being referred to is exactly one. The explanation he gives is of pragmatic nature: the singular is more informative and specific than the plural. Hence, in agreement with Grice's (1967) maxims of Quantity, Quality and Relevance, when given two expressions, one must pick the most specific one, assuming it is not more complex than the general expression. The singular form fits this description. First, because it specifically refers to one entity, unlike the plural, that can refer to one or more, and second, because it is morphologically simpler.

Let us now examine the functioning of both plural shapes in TA in light of the two competing views discussed above. The point here is first to test the use of plural shapes in questions and downward entailing environments and see if they behave in the same way as the English plural. The second purpose of this investigation is to find out whether the two plural shapes in TA behave similarly to one another with regards to number interpretation (or inclusivity versus exclusivity). Mathieu (2014) argues that although Arabic has two different plural forms, based on Saudi Arabic data, they semantically behave the same; they are both inclusive in the relevant contexts. Let us test this claim for TA. The examples in (9) illustrate the use of broken plurals in a question (9-a), a negation (9-b), and a conditional (9-c).

- (9) a. Famma ʃbebek fi biru-k? [Tunisian Arabic]
 there window.PL in office-POSS.2SG
 'Are there any windows in your office?'
 b. Ma rit-ef kleb f-el jarda.
 NEG saw.1SG-NEG dog.PL in-the garden.
 'I did not see any dogs in the garden.'

- c. Ken ʕand-ek wled jib-hom.
 if have.2G child.PL bring.2G-3.PL
 ‘If you have children, bring them.’

The broken plural shapes in (9) can all be interpreted inclusively. For instance, the question in (9-a) can be answered with ‘yes, one’, even though the question is whether the office has windows. ‘No, only one’ would not be an appropriate answer. (9-b) would be false if only one dog was seen. Finally, in the conditional in (9-c) (also a downward entailing context), it is implied that the person can bring any number of children, even one. I now turn to the interpretation of sound plurals in the same environments (10).

- (10) a. Famma muhands-een fi biru-k? [Tunisian Arabic]
 there engineer-MASC.PL in office-POSS.2SG
 ‘Are there any engineers in your office?’
- b. Ma rit-ef fannen-et f-el ḥaffa.
 NEG saw.1SG-NEG artist-FEM.PL at-the party.
 ‘I did not see any female artists at the party.’
- c. Ken t-aʕraf saḥafi-een jib-hom.
 if 2G-know journalist-MASC.PL bring.2G-3.PL
 ‘If you know journalists, bring them.’

The sound plural shapes in (10) can also be interpreted inclusively. For instance, the question in (10-a) can be answered with ‘yes, there is one’, even though the question is whether there are engineers. ‘No, only one’ would not be an appropriate answer. (10-b) would be false if only one female artist was seen was seen. Finally, in the conditional in (10-c) (also a downward entailing context), it is implied that

the person can bring any number of journalists, even one. I conclude from these tests that both plural shapes have the same interpretations in TA, regardless of the context.

Another claim by [Mathieu \(2013, 2014\)](#) is that the plural of the singulative in Arabic behaves differently from other plurals with regards to their interpretation. Unlike broken and sound plurals of count nouns, the plural of the singulative is odd in questions and downward entailing environments. Mathieu argues that the plural of the singulative is exclusive and paucal in all contexts. Paucity is a well-defined notion in Arabic. The paucal is used to refer to a small number of distinct entities, the equivalent of the English quantifier “a few” ([Corbett, 2000, 22](#)). While its use has no specific upper bound in English, languages that morphologically mark the paucal usually have variable cut-off points after which the plural is used ([Harbour, 2014](#)). In Arabic, the marker of paucity is used for persons and things that do not exceed ten in number ([Wright, 1933; Ojeda, 1992](#)). It turns out that this restrictive use of the plural of the singulative also applies in TA, where it is exclusive and can only refer to paucals. Consider the example in (11), illustrating the use of the plural of the singulative in different contexts.

- (11) a. El marʃi i-biʕ-u fiħ ħut/#ħut-a-at. [Tunisian Arabic]
 the market 3-sell-PL in.3SG fish.MASC.COLL/fish-FEM.SING-PL
 ‘In this market, they sell fish.’
- b. ʕand-ek ħut/#ħut-a-at?
 have-2SG fish.MASC.COLL/fish-FEM.SING-PL
 ‘Do you have fish?’
- c. ʃrit ħut/ħut-a-at.
 bought.1SG fish.MASC.COLL/fish-FEM.SING-PL

‘I bought fish.’

The examples in (11) reveal that while the collective shape (i.e., *ħut* ‘fish’) can be used in the same contexts as any plural shape and yield the same interpretation, this is not the case for the plural of the singulative. First, it is important to note that the infelicitous plurals in (11) are *feminine*. This provides strong evidence that they pluralize the singulative, and not the collective, which is always masculine. Although the concept of the plural of a unit created out of a collective seems redundant for the speaker of a non-singulative language, the result has a different denotation, as seen in (11). (11-a) illustrates a context referring to a large quantity of fish. While the collective form is suitable for this context, the plural of the singulative is not. This is because the plural of the singulative is a paucal and its use is limited to contexts referring to small quantities. Given that the plural of the singulative is restricted in its number reference, it is not expected to fit in questions and downward entailing environments where no specific number of entities is presupposed.

This prediction is borne out, as shown in (11-b). In the context where a client is asking the merchant if they sell fish, the use of the collective form would be more suitable. If the client used the plural of the singulative, it would mean that the client is looking to buy “more than one, but less than eleven” fish. This is very unlikely, since in general, the question would be about fish in general, as a kind/species, and not about a specific quantity. (11-c) illustrates a case where both forms can be used, each yielding a different interpretation. The use of the collective form would mean that the person bought fish, regardless of the quantity. It could be one, four or a hundred. The use of the plural of the singulative would mean that they bought a few

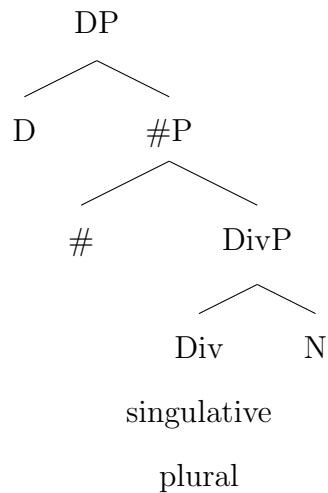
fish. Not one, not twenty, but some number between two and ten. Therefore, I use the term “paucal of the singulative” henceforth instead of “plural of the singulative”.

One important clarification is that I do not claim that the use of the paucal of the singulative is not possible in question or downward entailing environments. I rather suggest that these uses yield a very specific reading in such contexts. For instance, consider the example in (12), illustrating the use of the paucal of the singulative *toffeħaat* ‘apples’ in a question (12-a) and a conditional (12-b).

- (12) a. ʕand-ek toffeħ-a-at? [Tunisian Arabic]
have.2SG apple-FEM.SING-PL
‘Do you have apples?’
- b. Ken ʕand-ek toffeħ-a-at n-aʕml-u gattu.
if have-2SG apple-FEM.SING-PL 1-make-PL cake
‘If you have apples, we can make a cake.’

The person asking the question in (12-a) is specifically asking about a few apples. If the interlocutor only has one apple, they will most probably answer “I only have one (not a few)”. If the collective shape was used in the question, the answer would rather be “yes, I have one”. The conditional in (12-b) implies that it takes a few apples to make a cake and that one would not be enough. If the interlocutor has one apple, they would most probably answer “I only have one, unfortunately (so we cannot make a cake)”. The use of the paucal of the singulative in the conditional also implies that there only needs to be a few apples to make the cake. If the collective form was used, then any quantity of apples would qualify to yield a positive answer, even half an apple.

(14)



To reconcile the data with the theory, [Ouwayda \(2014\)](#) first argues that the plural marking that co-occurs with the singulative is not syntactically or semantically the same as other plural marking in the language. Despite appearances, the singulative and the feminine sound plural morpheme can coexist, and they do not compete for the same slot. In terms of their distribution, sound plurals of the singulative show three syntactic distinctions in Lebanese Arabic. First, unlike plural marked nouns in Arabic, nouns in which plural marking and the morphological classifier (singulative) *-a* co-occur cannot occur bare. This is shown in (15).

- (15) *ʃeft samk-e-et bi-l-baħer. [Lebanese Arabic]
saw.1SG fish-SING-FEM.PL in-the-sea
'I saw fish in the sea.'

Second, Arabic plural nouns are typically licit following quantifiers, but nouns in which plural marking and the singulative co-occur cannot occur following quantifiers, as shown in (16).

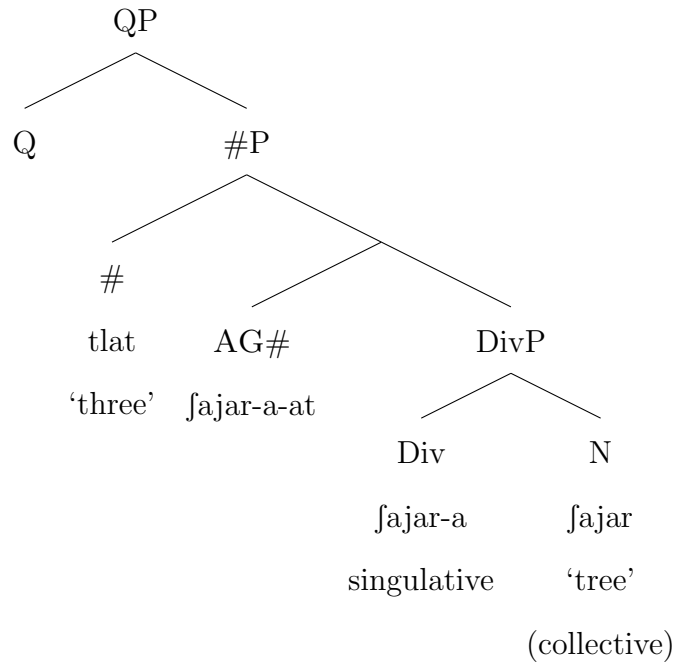
- (16) **ʃeft ktiir samk-e-et bi-l-baħer.* [Lebanese Arabic]
 saw.1SG many fish-SING-FEM.PL in-the-sea
 ‘I saw many fish in the sea.’

Finally, unlike other plural nouns in Arabic, nouns in which plural marking and the morphological classifier *-a* co-occur require the presence of a definite determiner or cardinal numeral in Lebanese Arabic, as illustrated in (17).

- (17) a. *S-samk-e-et muʃ hon.* [Lebanese Arabic]
 the-fish-fem.sing-pl not here
 ‘The fish are not here.’
- b. *ʃeft tesʃ samk-e-et bi-l-jaat*
 saw.1SG nine fish-SING-FEM.PL in-the-sea
 ‘I saw nine fish in the sea.’

Based on the syntactic constraints of plural marked nouns containing the singulative *-a*, [Ouwayda \(2014\)](#) proposes that this specific type of plural marking is an agreement marker that is dependent on the presence of a cardinal numeral or numeral-like element in # (henceforth AG#). She proposes that since cardinal numerals and definiteness have different licensing abilities for the agreement marker which co-occurs with the singulative, cardinal numerals and quantifiers are syntactically distinct. She therefore proposes the augmentation of the structure to allow cardinal numerals and quantifiers to merge in different locations. Specifically, she proposes a DP structure containing both a merger site for quantifiers Q, and a separate projection where cardinal numerals, specifying cardinality (#). The structure is schematized in (18).

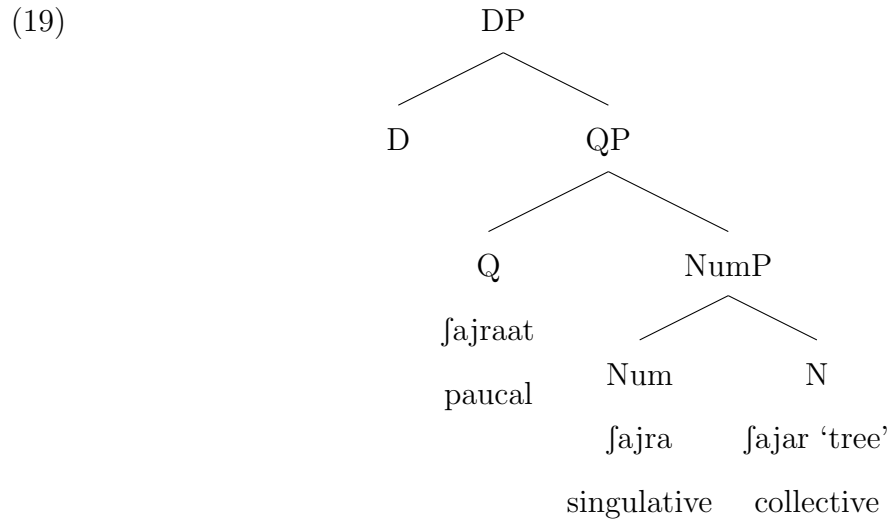
(18)



In the augmented structure proposed in (18), we can see that *ʃajaraat* ‘trees’, which appears to be a plural form, is an agreement marker (AG#), licensed by the presence of a cardinal (or determiner) in #. In this view, cardinals are distinct from quantifiers, which are hosted in Q. This structure reflects the data from Lebanese Arabic, where “plural” marked singulatives can occur with cardinals and definite determiners, but not with quantifiers.

The data from TA exhibit a different behaviour. First, the paucal of the singulative can appear with quantifiers, definite determiners and cardinals alike. They can also occur bare. More importantly, they have a paucal denotation, which is not present in Lebanese Arabic. For these reasons, and because they express a notion of quantity, I argue that they are actually paucals, thus, hosted in Q (in line with

Mathieu’s (2009; 2013) proposal for Saudi Arabic), as illustrated in (19).



I make no distinction between cardinals and quantifiers, since they seem to have the same distribution in TA, that is, they can both combine with the same types of nouns. Therefore, I do not distinguish them in the syntactic structure. If the paucal of the singulative combines with a numeral or quantifier, the latter can further be adjoined to QP. Although my data and analysis differ from Ouwayda’s (also see [Borer and Ouwayda 2010](#)), our common observation is that the sound plural morpheme takes special meaning only when co-occurring with the singulative. This is a puzzle that I explore further throughout this dissertation.

I have shown in this section that both plural shapes are used in the same environments and yield the same semantic interpretations in TA. They are both inclusive in questions and downward entailing environments, and they are unrestricted in terms of number reference. However, their use contrasts with that of the paucal of the sin-

gulative. As I have shown in the last set of examples and following [Mathieu \(2009, 2014\)](#), the paucal of the singulative is exclusive and paucal in all contexts.

4.2.3 Summary

In this section, I discussed the nature of the broken plural, with regards to its morphological and semantic status. I described two different views of broken plural formation: one treats them like stem-forming processed/Level I derivation ([Ratcliffe, 1998](#); [Kihm, 2003](#); [Acquaviva, 2008](#); [Lahrouchi and Lampitelli, 2014](#)), and the other treats them like inflectional forms derived from a singular ([Hammond, 1988](#); [McCarthy and Prince, 1990a,b](#)). Focusing on their semantics and contexts of use, I showed that broken plurals are equivalent of sound plurals, and concluded that both shapes are inflectional.

I also contrasted the use of both plural shapes in TA with that of the plural of the singulative. Following [Mathieu \(2009, 2014\)](#), I showed that the latter is an exclusive plural and defined them as paucals.

4.3 Plurals in a contrastive use

[Boudelaa and Gaskell \(2002\)](#) claim that for almost every broken plural shape, there is also a sound plural shape available. I address this claim by showing that plural shapes can alternate in certain contexts in TA. I argue that these alternations are determined by semantics and that they only occur in contexts where the broken plural is the default but the speaker uses the sound plural to convey a very specific

meaning (i.e., a paucal). I show that the sound plural in these contexts has a denotation similar to that of the paucal of the singulative, and consequently propose that this results from a reclassification of a count noun into the collective system.

4.3.1 Data

The discussion about Arabic plurals carried out in the previous section exposed the traditional view that each noun is assigned one shape or the other. Although analyses are divided on the nature of each plural form, there is a widespread view that the choice of a plural shape is based on the canonicity of the root. Canonical roots take the broken plural, and uncanonical stems are pluralized by suffixation (McCarthy and Prince, 1990a,b). To the best of my knowledge, no principles other than morphological/lexical ones have been claimed to govern the choice of a plural shape in Arabic.⁴

In this section, I present data from TA suggesting that morphological/lexical identity does not suffice to justify the choice of a plural shape. Indeed, it is possible for some nouns in TA to take both plural shapes, which is unexpected if we consider that the choice of a plural is determined by the morphological shape of the stem. This leads me to further explore the contexts of use of each plural and look for the underlying principle conditioning the choice of one shape over the other. Let me first illustrate with the example of the noun *meʕza* ‘goat’ that can either be pluralized by stem change (20-a) or suffixation of the feminine plural marker *-et* (20-b).

⁴One notable exception is Lahrouchi and Ridouane’s (2016) paper about the semantic contrast between plural forms in Moroccan Arabic. I argue for a similar contrast but based on different principles.

- (20) a. tfoḷ i-rabbi f-el mʕiiz [Tunisian Arabic]
 boy 3.MASC.SG-breed in-the goat.PL
 ‘a boy who breeds goats’
- b. Ken fallaḥ ʕandu arð w ʕwayya meʕz-et.
 was farmer have-3.MASC.SG land and few goat-FEM.PL
 ‘He was a farmer who owned a land and a few goats.’

I argue that when a single noun displays both plural forms, this results in a contrast in meaning. The proposed hypothesis is that the notion of contrast is essential to the interpretation of TA plurals. The sound plural behaves exactly like the broken plural when assigned in a lexically-based fashion. It is only in alternation with the broken plural that the sound plural gives rise to special meaning. More precisely, the sound plural has an exclusive and paucal reading when assigned to nouns that usually take a broken plural. In these cases, the use of the sound plural is semantically motivated and this particular plural serves a specific interpretive function that contrasts with the semantics of the broken plural, which has an inclusive semantics.

Let us begin by defining the context of alternation between plural shapes. To begin, an important clarification is that I do not claim that there is a systematic semantic difference between broken plurals and sound plurals in *all* contexts. I only consider cases where the broken plural could have been used (e.g., canonical stem), but the sound plural was used instead, and I investigate the semantic motivation behind this choice. For the sake of clarity, I refer to this specific use of the sound plural as *the contrastive sound plural* or *the sound plural of canonical stems*. I show that the choice of a plural in variable contexts is based on number sensitivity. I list

the nouns to be used in the following examples in Table 4.1 below. All the nouns in Table 4.1 have canonical shapes and consequently take a broken plural. However, they can also take a sound plural in specific contexts.

Table 4.1: Tunisian Arabic canonical nouns and their plurals

SINGULAR NOUN	DEFAULT PLURAL (BROKEN)	SPECIAL PLURAL (SOUND)
<i>meʕza</i> ‘goat’	<i>mʕiiz</i> ‘goats’	<i>meʕz-et</i> ‘goats’
<i>saḥfa</i> ‘bowl’	<i>shaaf</i> ‘bowls’	<i>saḥf-et</i> ‘bowls’
<i>filem</i> ‘movie’	<i>aflem</i> ‘movies’	<i>film-et</i> ‘movies’
<i>bseklet</i> ‘bike’	<i>bsekel</i> ‘bikes’	<i>bseklet-et</i> ‘bikes’
<i>korrassa</i> ‘notebook’	<i>krarass</i> ‘notebooks’	<i>korrassa-at</i> ‘notebooks’

Consider the attested⁵ examples in (21), where canonical nouns unexpectedly take the sound plural.

- (21) a. L-ʔab ken fallah ʕand-u arð w ʃwayya meʕz-et. [Tun. Arabic]
 DEF-father was farmer have-he land and few goat-FEM.PL
 ‘The father was a farmer who had a land and a few goats.’
- b. Kel saḥf-et maḥleh-om.
 these bowl-FEM.PL beautiful-they
 ‘These bowls, they’re pretty.’
- c. Tfarr-ejt fi film-et nhar sebt.
 watch-I in movies-FEM.PL day Saturday
 ‘I watched some movies on Saturday.’

As a preliminary observation, all examples in (21) show contexts where the use of a paucal form is suitable. For instance, note the presence of the quantifier *ʃwayya* ‘a few’ in (21-a) with the use of the sound plural *meʕzet* ‘goats’. In (21-b), the

⁵All the examples in this section were collected from the online Tunisian Arabic corpus (<http://www.tunisiya.org/>) or were provided by native speaker consultants.

speaker⁶ was specifically referring to two bowls. After being asked why she did not use the prescribed broken plural form, *shāf*, she replied “because there are only two of them”. Finally, in (21-c), the speaker was referring to the movies he watched on Saturday (unlikely to exceed ten). This sentence can be translated as “I watched some/a few movies.” All the sound plurals in (21) have paucal reference and cannot refer to more than ten/a few entities. The prediction is that these shapes would not be suitable in contexts where reference is made to more than a few entities. This is illustrated in (22).

- (22) a. #tfolli-rabbi f-el meʕz-et [Tunisian Arabic]
 boy he-breed in-the goat-FEM.PL
 ‘boy who breeds goats’
- b. #ħall-et ħanout ma t-bif ken es-saħf-et.
 open-3.FEM.SG store NEG 3.FEM.SG-sell only the-bowl-FEM.PL
 ‘A new store opened that only sells bowls.’
- c. #L-ʕaam hekk-a kol jemʕa n-emʕi l-cinema tfarr-ejt fi
 DEF-year that-3.masc.sg each week 1.SG-go to-movies watch-1.SG in
 barʕa film-et.
 many movie-fem.pl
 ‘That year I went to the movies every week. I watched a lot of movies.’

The examples in (22) show that the use of sound plurals of canonical nouns results in paucals. For instance, in the context of (22-a), the boy makes a living out of goat breeding, so the speaker is clearly referring to more than a few goats, and consequently, the sound plural is unsuited for this context. Similarly, (22-b) refers to more than a few bowls, as a store specialized in bowls would unlikely be selling less

⁶Personal communication.

than ten bowls. In (22-c), reference is made to more than a few movies, considering the speaker wants to express that they watched a lot of movies over the year. This explains why the sound plural cannot be used in the examples in (22). On the other hand, broken plurals of such nouns are not number sensitive and can be used in all the contexts presented above. Originally, the examples in (22) all came from attested examples where the broken plural was used, as seen in (23).

- (23) a. tfoll i-rabbi f-el mfiiz [Tunisian Arabic]
 boy he-breed in-the goat.PL
 ‘boy who breeds goats’
- b. ħall-et ħanout ma t-biʃ ken es-šhaf.
 open-she store NEG she-sell only the-bowl.PL
 ‘A new store opened that only sells bowls.’
- c. L-ʕam hekk-a kol jemʕa n-emʃi l-cinema tfarr-ejt fi barʕa
 DEF-year that-he each week I-go to-movies watch-I in many
 aflem.
 movie.PL
 ‘That year I went to the movies every week. I watched a lot of movies.’

Not only can the broken plural refer to numbers higher than ten, but they can also apply to all cases where the paucal is felicitous. Consider the examples in (24), which show that broken plurals can be used in paucal contexts.

- (24) a. L-ʕab ken fallaħ ʕand-u arð w chwayya mfiiz.[Tunisian Arabic]
 DEF-father was farmer have-he land and few goat.PL
 ‘The father was a farmer who had a land and a few goats.’
- b. Kel šhaf maħleh-om.
 those few bowl.PL beautiful-they
 ‘These (few) bowls, they’re pretty.’

- c. Tfarr-ejt fi affem.
watch-I in movies.PL
'I watched some/a few movies.'

The mapping between form and meaning is not exactly one-to-one, as the use of the broken plural is not strictly associated to with non-paucal reading. Instead, as seen in (24), the broken plural can be used in all contexts, including the paucal. The general idea is that the broken plural, in addition to being the default, morphologically unmarked plural, is also semantically unmarked, while the sound plural carries special meaning.

Lahrouchi and Ridouane (2016) also observed that a single noun can display both plural forms in Moroccan Arabic. For instance, the Moroccan Arabic noun *taswera* 'photograph' has two possible plural forms: the broken plural *tsawer* and the sound plural *teswerat*. Lahrouchi and Ridouane also noted the semantic contrast between the two forms: the sound plural indicates a definite number, while the broken plural has a collective reading. They account for this contrast by positing that broken plurals and sound plurals reside in different syntactic locations, respectively the *n* projection, and the standard Num projection. My account differs from Lahrouchi and Ridouane's. First, while they make a systematic distinction between the two plural shapes, I only consider the *contrastive* sound plural to be different from the broken plural. Sound plurals of uncanonical nouns show no semantic contrast with broken plurals. Moreover, although I also associate the plural with a collective interpretation to be located lower than the plural with a paucal interpretation, I consider the broken plural to be the result of a productive operation and thus to

be higher than the *n* projection, as I have previously argued. I also draw a parallel between the paucal reading of the sound plural used in contrast and that of the paucal of the singulative, suggesting that the canonical noun pluralized with a sound plural was in fact reclassified as a noun belonging to the collective paradigm.

4.3.2 Parallels with the collective system

In [Section 4.2.3](#), I discussed the denotation of the paucal of the singulative. I showed that, unlike other plural shapes and the collective base, the paucal of the singulative is more restricted in terms of its contexts of use. While plural nouns of the count class and the morphologically unmarked collective have general use (can refer to any number, including the singular in certain contexts), the plural marker affixed to a singulative noun results in a paucal. This prevents them from having inclusive reference in questions and downward entailing contexts. Based on the data presented in the previous section, it appears that the sound plurals of canonically-shaped nouns have the same interpretation as the paucal of the singulative. Moreover, they are also typically not used in questions, unless it implies a specific number of entities. To illustrate, consider the example in (25), where the two possible plural shapes of the noun *besklet* ‘bike’ (canonical stem) are shown.

- (25) Tsallh-u l-bsekel/#besklet-et? [Tunisian Arabic]
 repair-2PL the-bike.PL/bike-FEM.SG
 ‘Do you repair bikes?’

In (25), in the context of a question, the sound plural *beskletet* is odd since it would

imply that the speaker is asking if they repair “more than one, but less than eleven” bikes. However, the use of the broken plural in this case refers to bikes in general, with no number presupposition. This shows that, like the paucal of the singulative, the sound plural of canonical nouns is not only paucal, but also exclusive in all environments.

Another interesting observation that relates the contrastive sound plural use to the paucal of the singulative, is that both shapes are feminine. The sound plurals of canonical Arabic nouns (even the masculine ones) are always feminine. In other words, only the sound plurals ending with the feminine plural marker *-at/-et* can compete with broken plurals in TA. Take for example the masculine noun *filem* ‘movie’ in (26). This noun has a canonical shape and can thus take a broken plural. It can also be pluralized by suffixation of the feminine plural marker *-et*, although its singular counterpart triggers masculine agreement on the adjective, as illustrated in (26-a).

- (26) a. T-farrej-t fi filem hlow. [Tunisian Arabic]
 1SG-watched-PERF in movie.SG beautiful.MASC.SG
 ‘I watched a beautiful movie.’
- b. aflem film-et *film-een
 movie.PL movie-FEM.PL movie-MASC.PL
 ‘movies, movies, movies’

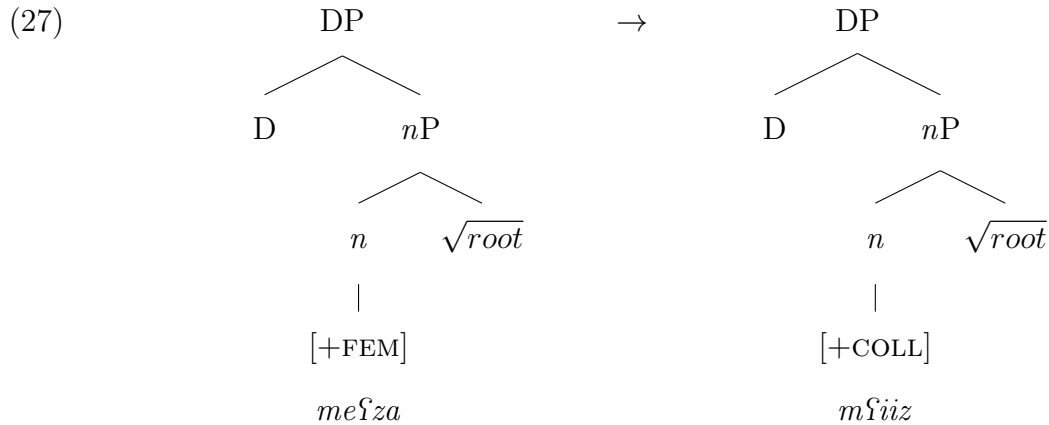
As shown in (26-b), the masculine sound plural is not an option for canonical stems, even when the nouns is masculine in the singular. This constraint on the gender of the contrastive sound plural raises a crucial point, given the important role of the

grammatical feminine gender in Arabic, and its close interaction with number. The paucal of the singulative is subject to the same constraint, and this is due to the fact that it pluralizes a feminine form that fills a dividing function. In the next section, I explore the possibility that the contrastive use of sound plurals is the result of a reclassification of a count noun into the collective system.

4.3.3 Analysis: a reclassification of the noun

According to (Wurzel, 1987, 87), one of the characteristics of a stable inflectional class is productivity. The integration of borrowings into the inflectional system, as well as transferals from other classes are tokens of productivity (see also Stolz 2001). In Chapter 3, I argued that the Arabic collective class is an inflectional one, providing arguments such as productivity and class stability. If Wurzel's predictions are correct, we should expect transferals between classes to be possible in this system. In what follows, I argue that the sound plurals of canonical nouns are indeed instances of transferals of count nouns into the collective class. The main arguments for this analysis are the use of the feminine sound plural ending and the paucal interpretation associated to them.

The proposal is as follows. A given noun (e.g., *meʕza* 'goat'), which is classified as a count noun, is reclassified as a collective one by the speaker, in order to convey a specific paucal meaning. This is done simply by marking *n* with the +[COLL] feature, as illustrated in (27).

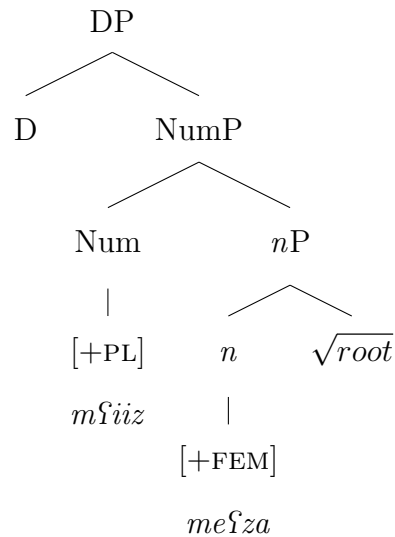


Note that the broken plural shape acts like the collective base in such scenarios. First, as I have previously shown, the broken plural and the collective are similar in their denotations. When in contrast with the sound plural or the paucal of the singulative, they are both interpreted as general, inclusive plurals. As further discussed in [Section 4.4](#), the broken plural is often viewed as a singular collective noun. This is evidenced by data where broken plural subjects trigger singular agreement on verbs and adjectives. This similarity between broken plurals and collective nouns was also discussed by ([Wright, 1933](#), 233). He notes that “[broken plurals] are, strictly speaking, *singulars* with a *collective* signification, and often approach in their nature to *abstract nouns*”. Therefore, considering this parallel between broken plurals and collectives, I propose that in the cases where a count noun is reclassified as a noun of the collective class, the broken plural is the equivalent of the collective, the singular form is the equivalent of the singulative, and the feminine sound plural is the equivalent of the paucal of the singulative.

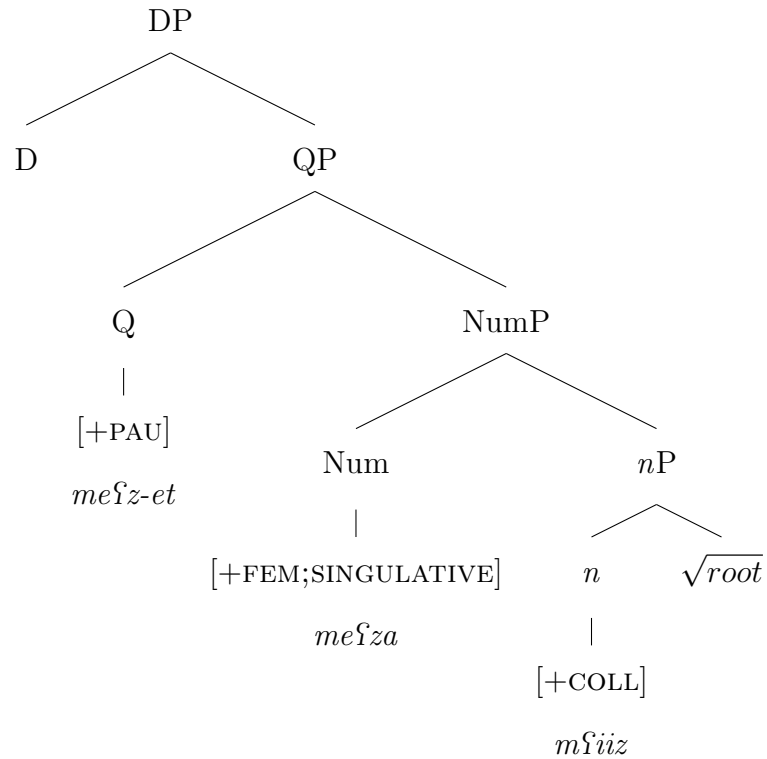
The effects of this change of features on *n* are the following. First, the gender

features are now interpreted on Num instead of *n* (remember that biological gender contrasts are neutralized on *n* for collective nouns). This explains why all canonical nouns pluralized by suffixation can only take the feminine sound plural ending. The feminine in this case is used as a singulative marker on *n*. A second effect of reclassifying a count noun into a collective one is that its pluralization (which is in fact a pluralization of the singulative) results in a paucal reading that is interpreted on Q rather than Num. This is evidenced by the fact that Num is already occupied by the singulative, and also by the paucal reading that results from the suffixation of *-at*. To illustrate, while the broken plural of a canonical noun in TA yields (28), the sound plural of such a noun results in (29).

(28) Broken plural



(29) Sound plural



One question that may arise from this proposal is the following. If a count noun such as *meɣza* ‘goat’ is transferred to the collective system, we expect instances of the singulative shape of this noun (without pluralization). My view on this matter is that the transferral into the collective class is only to fill a specific purpose, that is, the use of the paucal. By “mimicking” the paucal of the singulative, the speaker can obtain a paucal reading for any noun. Outside of this specific use and context, the speaker does not need to transfer a count noun to the collective system, since the available morphological tools are enough to express both individual and sums reference, without resorting to noun reclassification.

4.3.4 Summary

In this section, I showed that the sound plural can be used on canonical nouns in TA, and that this use is conditioned by semantic factors. While the use of the broken plural results in a general and inclusive reading, the sound plural is in fact better described as a paucal in this case. When the feminine plural marker *-at* is affixed to a canonical noun, we obtain a paucal and exclusive reading, similar to the paucal of the singulative. I therefore proposed that the use of the sound plural on canonical nouns to serve a specific semantic purpose is done by reclassifying the count noun into a collective noun. This reclassification has the effect of changing the locus of interpretation of both the gender and number features. The gender feature is interpreted on Num (instead of *n*), and the number feature is interpreted on Q (instead of Num). This possibility of transferring a noun from one class to another provides an additional argument to the claim that the collective is an inflectional class Wurzel (1987); Stolz (2001).

4.4 Broken plurals with a group reading and pluratives

I have previously shown, based on work by Hasselbach (2014a,b), that what is now known as the feminine suffix *-a* has not always been correlated to the feminine gender. Reconstruction studies of Proto-Semitic indicate that the role of *-a* followed a sequence of historical changes, including nominalization, singulative marking, group formation, and finally gender marking. In Chapter 2, I discussed how gender in Ara-

bic is involved in nominalization, and further, how it can be interpreted as referring to biological sex. Both these functions result from the gender feature being hosted on *n*. Then, in [Chapter 3](#), I discussed the case of the singulative, and argued that this operation results from gender being interpreted on Num. I now turn to the last scenario, where the feminine suffix expresses group formation. As with other roles played by *-a*, the synchronic facts again seem to show some traces of the function of group formation observed in Proto-Semitic.

4.4.1 Broken plurals with a group reading

An unusual agreement pattern can arise with broken plurals in TA. A verb may agree with a broken plural subject in all ϕ -features (30-a) or, unexpectedly, in feminine singular (30-b), even when the subject refers to masculine entities.

- (30) a. El rjel xerj-u. [Tunisian Arabic]
 the man.MASC.PL went.out-3.MASC.PL
 ‘The men went out.’
- b. El rjel xerj-et.
 the man.MASC.PL went.out-3.FEM.SG
 ‘The men went out.’

Note that this agreement mismatch (I am using “mismatch” as a pre-theoretical term here) not only involves number, but also gender. This phenomenon has been discussed in previous literature for Standard Arabic and modern spoken dialects ([Wright, 1933](#); [Belnap, 1991](#); [Brustad, 2000](#); [Zabbal, 2002](#); [Dali and Mathieu, 2016, 2020](#)). What has been observed in all these accounts is that the masculine plural

agreement correlates with a distributive interpretation, whereas the feminine singular agreement correlates with a collective reading. The same facts obtain in TA. In the right contexts, (30-a) means that the men went out separately, one by one, but (30-b) means that they went out together, as a group. Only broken plurals are subject to this phenomenon, as evidenced by the ungrammaticality of (31-b), where the subject is a masculine sound plural, which fails to trigger feminine singular agreement on the verb.

- (31) a. *El muʕalm-een rajʕ-et l-el-biru. [Tunisian Arabic]
the teacher-MASC.PL returned-3.FEM.SG to-the-office
‘The teachers went back to the office.’
- b. El muʕalm-een rajʕ-u l-el-biru.
the teacher-MASC.PL returned-3.MASC.PL to-the-office
‘The teachers went back to the office.’

It is important not to confuse this pattern with other seemingly related phenomena in Standard Arabic, namely partial agreement and deflected agreement. Partial agreement is a well-known phenomenon in Standard Arabic, whereby the verb in VSO sentences is inflected in the singular even though the subject noun is plural. This is illustrated in (32). SVO sentences in Standard Arabic trigger all the ϕ -features of the verb (e.g., 3, MASC, PL), as shown in (32-a). On the other hand, in VSO sentences, as in (32-b), the verb is obligatorily inflected in the singular even though the subject noun is plural.

- (32) a. Al-ʔawlaad-u jaaʔ-uu. [Standard Arabic]
the-boys-NOM came-3.MASC.PL
‘The boys came.’

- (34) *ja:ʔa-ti al-ʔawlaad-u. [Standard Arabic]
 came-3.MASC.SG the-boys.NOM
 ‘The boys came.’

In TA, on the other hand, verbs agreeing with non-humans and inanimate subjects can also inflect in the third person feminine singular, but it is an optional process, as seen in (35). More importantly, it is also possible with human subject nouns, as in (30).

- (35) a. El biben tsakkr-u. [Tunisian Arabic]
 the door.PL closed-MASC.PL
 ‘The doors closed.’
 b. El biben tsakkr-et.
 the door.PL closed-FEM.SG
 ‘The doors closed.’

One last observation is that the feminine singular agreement appears not only on the verb, but also on adjectives and other such categories. In (36), the agreement on the quantifier *kolli* ‘all’ and the adverbial expression *mʕa bʕaḏha* ‘together’ are inflected in the feminine singular. This indicates that agreement is controlled by the subject noun, which is, as I argue in view of this and other evidence, syntactically feminine and singular.

- (36) El wled el kolli xerj-et mʕa bʕaḏ-ha. [Tun. Arabic]
 the boy.MASC.PL the all.FEM.SG left-3.FEM.SG with other-3.FEM.SG
 ‘The boys all left together.’

The following two examples provide further evidence that feminine/singular agree-

ment is continuous in the cases at hand.

- (37) ʔand-ek meʔekel nafsiy-ya. [Tunisian Arabic]
have-2SG problem.BPL personal-FEM.SG
'You have personal problems.'
- (38) xlaʔt el beb beʔ taʔmel el ʃweh heð-i lkol?
forced.2.SG the door will do.IMP.2.SG the scandal.BPL this-FEM.SG all
'You forced the door to make all these scandals?'

I have presented the case of an unusual agreement pattern triggered by broken plurals in TA and possibly other dialects. What is important to retain from this phenomenon is that it is optional, concerns broken plurals, and it involves both gender and number features. I now turn to a discussion about pluratives, which appear to present similar agreement patterns.

4.4.2 Pluratives

Pluratives, briefly discussed in Section 3.2.3, are Arabic singular nouns referring to ethnic groups or members of a profession (Fassi Fehri, 1984, 1988, 2012, 2018b). The plurative designates a group, and is morphologically characterized by the feminine suffix *-a*, in addition to triggering feminine singular agreement on targets (verbs or adjectives). Again, the result is usually interpreted as an integrated whole, as opposed to a distributive interpretation (Fassi Fehri, 2012, 2018b). Examples of pluratives are provided in (39).

- (39) a. *saydal-ii* ‘pharmacist-masc’ → *sayaadil-a(t)* ‘the corps of pharmacists’
 [Standard Arabic (Fassi Fehri, 2018b, 10)]
- b. *xayyat* ‘tailor’ → *xayyat-a* ‘tailor-fem’, ‘tailors’ (the corps of tailors)
 [Moroccan Arabic (Fassi Fehri 2018:10)]
- (40) a. *ʔafriq-ii* ‘african-masc’ *ʔafaariq-a(t)* ‘Africans’
 [Standard Arabic (Fassi Fehri, 2018b, 10)]
- b. *jabl-ii* ‘inhabitant of the mountain-masc’ *jbal-a* ‘inhabitants of the mountain’
 [Moroccan Arabic (Fassi Fehri, 2018b, 11)]

An interesting point regarding this data is that the feminine singular marker *-a* can also be used to denote feminine members of mating pairs. As an illustration, consider the TA examples in (41) and (42).

- (41) a. *ħajjem* *ħajjem-a*
 hairdresser.MASC.SG hairdresser-FEM.SG
 ‘male hairdresser, female hairdresser or a group of hairdressers’
- b. *bedwi* *bedwi-a*
 bedouin.MASC.SG bedouin-FEM.SG
 ‘male bedouin, female bedouin or a group of bedouins’

The feminine singular nouns in (41) have ambiguous meaning. They can either refer to female individuals or to a group of individuals (of either biological gender). The *-a* morpheme that appears on the noun can be interpreted either as a marker of feminine biological gender, or as a group marker. Verbal agreement can help disambiguate

4.4.3 Analysis: broken plurals are hybrid nouns

Wright’s (1933) original insights about the broken plural are summarized in the following quote:

- (44) “As regards their meaning the *plurales fracti* [broken plurals] differ entirely from the sound plurals; for the latter denote several *distinct* individuals of a genus the former a number of individuals viewed *collectively*, the idea of individuality being wholly suppressed. [...] The *plurales fracti* are consequently, strictly speaking, *singulars* with a *collective* signification, and often approach in their nature to *abstract nouns*. Hence, too, they are all of the *feminine* gender, and can be used as masculine only by a constructio ad sensum.” (Wright, 1933, 233)

This quote brings out many important characteristics of the broken plural. First, it highlights the idea that the broken plural can have this singular collective effect, and crucially, associates this interpretation with the feminine gender. This quote also suggests that while broken plurals are syntactically feminine, the masculine can be used as a semantic agreement to avoid ambiguity.

To account for the alternation in agreement patterns, Dali and Mathieu (2020) propose that broken plurals are hybrid nouns, which have mismatching syntactic/semantic ϕ -features and are thus able to trigger either syntactic or semantic agreement (Landau, 2015; Corbett, 2000, 2015; den Dikken, 2001; Wechsler and Zlatić, 2003; Danon, 2011, 2013; Smith, 2015). To illustrate, let us take the example

of group nouns in British and Canadian English (*committee, team*). These have been argued to be hybrid nouns (Corbett, 2000, 2015; Wechsler and Zlatić, 2003; Smith, 2015). Consider the examples in (45). They show that it is possible for the verb to be inflected in the singular (45-a) or in the plural (45-b).

- (45) a. The committee *is* meeting next week.
 b. The committee *are* meeting next week.

The idea behind treating group nouns as hybrids is to suggest that they have the feature matrix in (46) where either the syntactic features or the semantic features can be accessed. In (45-a), the syntactic features are activated whereas in (45-b), the semantic features are at play.

$$(46) \quad \left[\begin{array}{ll} \text{syn: 3} & \text{SG} \\ \text{sem: 3} & \text{PL} \end{array} \right]$$

This contrasts with plural subjects, since they have the feature grid in (47).

$$(47) \quad \left[\begin{array}{ll} \text{syn: 3} & \text{PL} \\ \text{sem: 3} & \text{PL} \end{array} \right]$$

Turning now to broken plurals in TA, the idea in Dali and Mathieu (2020), is that broken plurals behave like British/Canadian English group nouns rather than plural subjects. Broken plural subjects in TA can agree in the singular or in the plural quite freely: with distributive (48), collective (49), as well as ambiguous predicates

(50).

- (48) a. Rjel i-koḥ-u. (distributive)
man.PL 3.MASC-coughing-PL
'(Some) men are coughing.'
- b. Rjel t-koḥ. (collective)
man.PL 3.FEM-coughing.SG
'(Some) men are coughing.'
- (49) a. El rjel tamm-u. (distributive)
the man.PL gathered-3.MASC.PL
'The men gathered.'
- b. El rjel tamm-et. (collective)
the man.PL gathered-3.FEM.SG
'The men gathered.'
- (50) a. El rjel xerj-u. (distributive)
the man.PL went.out-3.MASC.PL
'The men went out.'
- b. El rjel xerj-et. (collective)
the man.PL went.out-3.FEM.SG
'The men went out.'

Dali and Mathieu (2020) propose that broken plurals are syntactically singular, and so they strictly agree with the verb or the adjective in the singular. Semantically, broken plurals are plural: when the agreement on the verb or the adjective is plural, it is the semantic features on Num that are accessed. Broken plurals also contain a feminine gender feature. This feature is the spell-out of a group reading. The grid for broken plurals is represented in (51).

$$(51) \quad \left[\begin{array}{l} \text{syn: 3 SG fem} \\ \text{sem: 3 PL fem/masc} \end{array} \right]$$

(52), on the other hand, gives the feature grid for sound plurals.

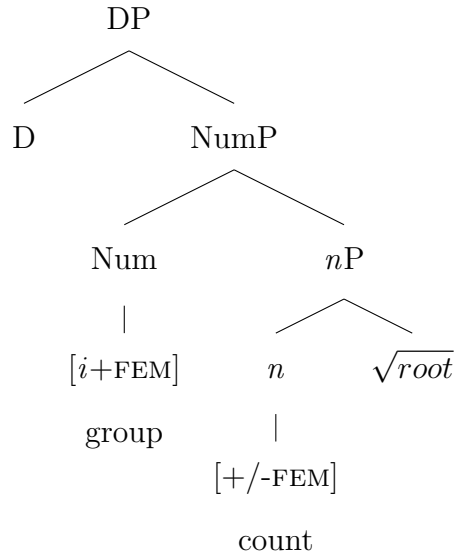
$$(52) \quad \left[\begin{array}{l} \text{syn: 3 PL fem/masc} \\ \text{sem: 3 PL fem/masc} \end{array} \right]$$

Based on the main proposal of this dissertation, I propose the following functioning for hybrid nouns. For broken plural nouns, the natural or arbitrary gender is on n , as is the case for all nouns of the count system. The gender feature on Num is i [+FEM], with an associated singular feature ([-PL])⁷ as seen in (53). If semantic features are accessed, agreement is based on the gender on n , and the number feature on n is [+PL].⁸ If syntactic features are accessed, then agreement is based on the gender on Num (feminine), and the number feature is singular by way of the feminine feature yielding a group reading.

⁷The correlation between interpretable feminine feature on Num and singular feature can easily be made by a co-occurrence rule like the one proposed in [Chapter 2](#) to account for the feminine gender or plural inanimates.

⁸The co-occurrence rule is no longer at play here, since the gender feature is on n , not Num.

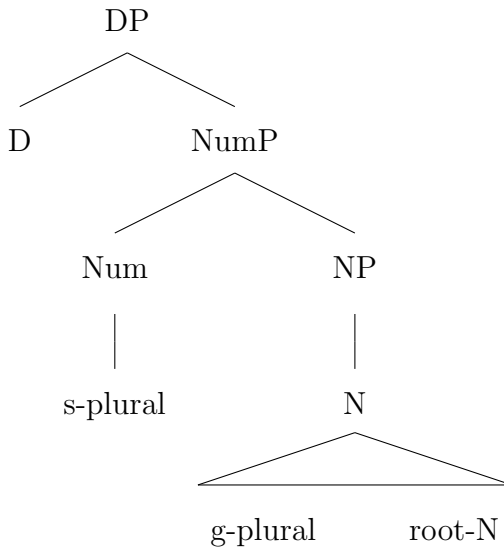
(53)



The idea that broken plurals are syntactically singular and feminine is found in (Wright, 1933, 233). In addition, there is diachronic evidence in favour of the view that broken plurals are singular. Historically, the singular was the only number in Semitic languages (Lipiński 2001: 242, Haelewyck 2016:153). Plural reference was expressed by the singular collective, now known as the broken plural shape. This form agreed in the singular only, as do collective nouns in Arabic. The suffixal (sound) plural is the result of later developments in the history of Semitic languages (Hasselbach, 2014a,b). This means that the broken plural was not syntactically plural in the old number system of Arabic and the plural is thus an innovation.

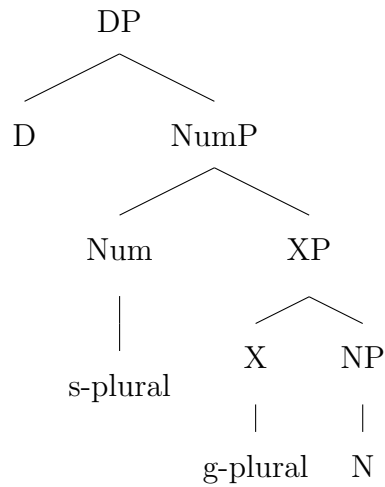
Another point to discuss is that group denotation is associated with the feminine in the count system (with broken plurals), and with the masculine in the collective system (with collective nouns). This difference resides in the fact that they do not belong to the same class. Gender features are interpreted differently, according to

(55)



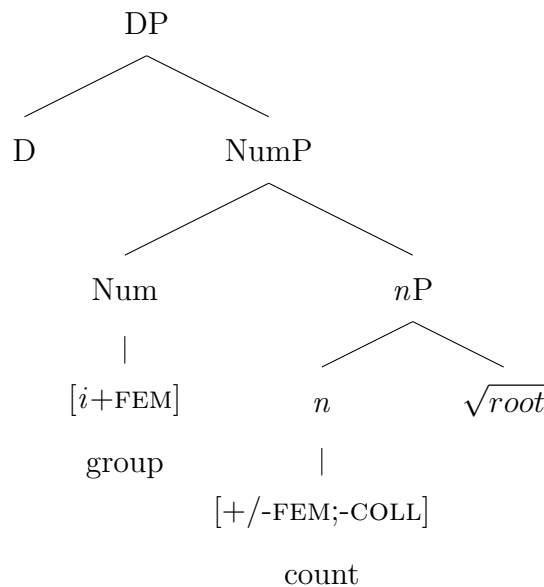
Zabbal (2002) later proposes another structure to account for the fact that g-plurals are productive (and therefore inflectional). He proposes, as before, that s-plurals are under Num, but g-plurals are under X (an undefined projection). A group operator is generated in X and it turns a plural NP into an atom. This is shown in (56).

(56)



shift triggering feminine singular agreement originates from the presence of a phonologically null *n* with a herd/clump meaning. This analysis is in line with Kramer’s (2015) original proposal that all gender features are hosted on *n*. The analysis also echoes other accounts of gender shifts on plural marking, dubbed “gender polarity”, as being generated on *n* due to the derivational nature of gender features (Lecarme, 2002; Acquaviva, 2008; Kramer, 2015; Lahrouchi and Lampitelli, 2014). My proposal also presents this phenomenon as a syntactic one, given the effect on semantic interpretation. However, based on the arguments presented throughout this dissertation in favour of a multi-level account of gender, I argue that the feminine is interpreted on Num in such cases.

(58)



The crucial point about these data and the analyses presented above is the fact that the gender feature—the feminine, more specifically—can be interpreted on Num and yield a group interpretation. This role of the feminine gender was suggested by

Hasselbach's (2014a; 2014b) research on the role of the feminine in diachrony, where group formation is enumerated as one of the roles.

4.4.4 Summary

In this section, I addressed the case of broken plurals and pluratives, which denote plural entities but can trigger feminine singular agreement on the verb and adjective. This type of agreement pattern shows a mismatch in terms of number, but also gender, since singular agreement is always associated to the feminine gender, even when the broken plural/plurative refers to masculine entities. Based on diachronic facts, suggesting that broken plurals were originally syntactically singular, and following a proposal by Dali and Mathieu (2020), I analyzed TA broken plurals as hybrid nouns. In this account, the broken plural is considered to be syntactically feminine/singular and semantically masculine/plural. Both sets of features can be accessed, depending on pragmatic factors.

4.5 Conclusion

In summary, this chapter was dedicated to the topic of broken plurals and their morphological formation and semantic interpretation.

I first introduced different views of the morphological nature of broken plurals, and concluded, with supporting evidence from TA data, that they have no special meaning and are, therefore, inflectional like sound plurals.

I then showed how sound plurals can be used in contrast with broken plural

shapes to convey a paucal interpretation, and proposed that this can be done by reclassifying count nouns into the collective class.

Finally, I showed how the feminine feature can be interpreted on the Num head to express group formation. This was supported by data from TA, where broken plurals can trigger feminine singular agreement on verbs and adjectives.

Chapter 5

A Case of Allosemy

5.1 Introduction

Throughout this thesis, a recurrent theme has been that the interpretation of the [+FEM] feature in TA depends on whether it is marked on a count or a collective noun. This is the consequence of the diachronic facts about gender in Semitic, and more specifically, the coexistence of two number systems, each making a distinct use of the feminine gender.

This chapter aims to formalize the generalization stated above, namely, that noun category determines the interpretation of gender in TA. Focusing on the feminine gender, I argue that the different meanings associated to [+FEM] are allosemic. I then extend this allosemic account to the multiple interpretations of the [+FEM, +PL] feature, phonologically realized as *-at*. This allows me to offer a simple formal account of a seemingly complex set of data. In sum, this proposal boils down to

claiming that a child only needs to learn the grammatical category of each noun, and that the interpretation of the feminine and plural morphemes simply follows from this knowledge.

The chapter is organized as follows. In [Section 5.2](#), I define allosemy and discuss potential pitfalls and advantages of such an account. In [Section 5.3](#), I discuss two cases of allosemy within the DP: one about proper names, and one about Num projections in Dagaare. Finally, I articulate a proposal about the feminine and plural features in TA in [Section 5.4](#), arguing that the different meanings of the feminine and the plural features are allosemic and are contingent upon the nature of the noun, which is marked on *n*. [Section 5.4](#) concludes.

5.2 What is Allosemy?

The term “allosemy” refers to the phenomenon in which a single morpheme can have multiple semantic realizations. This definition is reminiscent of that of allomorphy, where a single abstract morpheme can have multiple phonological realizations. Indeed, allosemy is most often discussed in the literature as a parallel of allomorphy in the LF domain ([Wood, 2012](#); [Marantz, 2013](#)). Due to this inevitable comparison and to lay the ground for a better understanding of allosemy, I begin by sketching an overview of allomorphy, a much more common and generally accepted concept. This gives us a better understanding of what happens on the semantics side of things.

Allomorphy is the phenomenon in which Vocabulary items, listed according to their morphosyntactic categories, compete to realize the resulting terminal elements

in Vocabulary Insertion. The criterion to determine the winner in such competition between VIs depends on the type of competition: it can be context-free or context-dependent (Halle and Marantz, 1993). In context-free insertion, the selected Vocabulary entry is basically the one whose features match the set of morphosyntactic features generated by the syntax and morphology on the terminal element. In this case, the competition is among different entries for the chance to spell out a particular set of features, and the entries differ only in the features they realize and in no other factor. The second type of competition, context-dependent insertion (or *conditioned allomorphy*), is more relevant to the present discussion. Like context-free Vocabulary Insertion, conditioned allomorphy also involves a choice among alternative Vocabulary Items. However, the choice in this case is not among items that differ in their substantive morphosyntactic features, but among items that differ in their insertion contexts and phonological features. The case of the English past tense suffix 0 is a well-known example. The latter is selected by *strong* verb stems (e.g., *put*, *read*), whereas *weak* verbs select the suffix /-t/ or /-d/ (e.g., *dwelt*, *played*) to realize the same set of tense features, as illustrated in (1). The substantive features ([+past], etc.) on the /-t/, /-d/, and 0 allomorphs are the same; they differ only in contextual features. In sum, conditioned allomorphy describes cases where allomorphy in a morpheme may be triggered by another morpheme (Bobaljik, 2000).

(1) PF instructions: phonological realizations of *Tns*

[+past] ↔ ∅ / [+strong] _____

[+past] ↔ /-t/ / [-strong]_____

[+past] ↔ /-d/

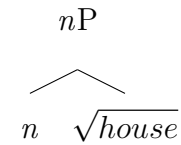
Assuming that morphemes are abstract before their spell-out at the PF and LF interfaces, it is natural to assume that morphemes can be as abstract with respects to semantics as they are with respect to phonology. Consequently, [Marantz \(2013\)](#) argues that there can be an equivalent of VI on the LF side that introduces semantic values to morphemes depending on the context. Let us call it “alloseme” or “Meaning insertion” to mark the parallel with Vocabulary insertion. If morphemes were semantically concrete in the syntax, we would identify them by their semantic content in the syntax rather than by their grammatical features. Having established this, allosemy is the phenomenon by virtue of which a given morpheme receives different semantic interpretations depending on its local syntactic context. While [Marantz \(2013\)](#) focuses on root allosemy (the choice of particular meanings for a root in a particular context), he also mentions the possibility for functional heads to be allosemic (see [Wood 2012](#); [Marantz 2013](#); [Wood and Marantz 2017](#)). [Harley \(2014\)](#) also supports the notion of allosemy in the domain of roots. She shows that roots are both phonologically and semantically underspecified and that their exponents and interpretations can be conditioned by the content and structure of their local syntactic environment. To illustrate, let us look at one case of root allosemy, followed by one case of functional head of allosemy.

[Marantz \(2013\)](#) illustrates root allosemy with the verb *to hou[z]e*, from the root HOUSE. In addition to showing contextual allomorphy (special voicing of the final fricative when merged with a little *v* head), this verb also shows contextual allosemy, since no literal *house* nor even a *container* is implied by the verb. In sum, the meaning of the root is conditioned by its syntactic environment, as illustrated in (2)

and (3).

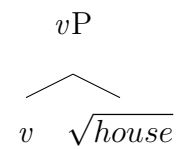
(2) *Root*

$[\sqrt{house}] \leftrightarrow \text{concept of a house}^1 / [n]$ _____



(3) *Root*

$[\sqrt{house}] \leftrightarrow \text{no concept of house} / [v]$ _____



Functional head allosemy can be illustrated by the different meanings of the *v* head (Marantz, 2009; Wood, 2012). Assuming two interpretations of the *v* head, a stative and a dynamic one (4), Wood (2012) has put forward an analysis where the interpretation of *v* is conditioned by the denotation of its complement (state or event).

(4) a. $[[v]] \leftrightarrow e_s.\text{activity}(e)$

b. $[[v]] \leftrightarrow e_s.\text{state}(e)$

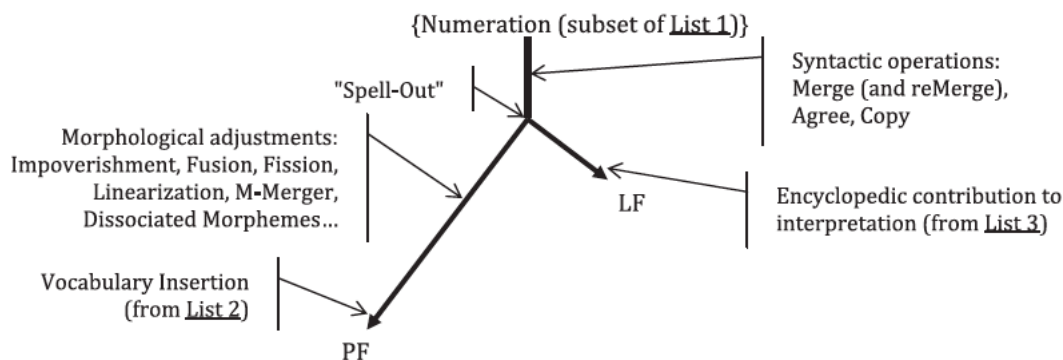
If the complement of *v* denotes a stative result (e.g., *opens the door*), then *v* will be interpreted as causative introducing an agent. Otherwise, it is interpreted as an activity (e.g., *runs*). These allosemes compete for the denotation of *v*, much like

¹Instead of providing a semantic denotation for the root, I assign it a vague notion of concept, since the root is not supposed to have fixed semantics before merging with the categorizing head, according to DM.

Vocabulary Insertion on the PF branch. The different semantic “flavours of *v*” are conditioned by the semantics of its sister.

To sum up, allosemy can be found both in roots and functional items, as long as their meanings are conditioned by the syntactic context in which they occur. This is the equivalent of contextual allomorphy, but on the LF branch, as illustrated in (5).

(5) The model: DM by [Halle and Marantz \(1993\)](#) (from [Harley 2014](#), 228)



5.2.1 Potential pitfalls of an allosemic account

The general idea behind contextual allosemy is that instead of endowing roots and functional heads with a semantics that remains constant across all its instantiations, we give each root or functional head like little *v* a certain number of possible semantic meanings that we will define as *allosemic*. This treatment allows syntax to be autonomous, like any root-based framework (e.g., Borer’s exo-skeletal model). As a side effect, this signifies that meaning can be potentially built up in different ways, and the same structure can have different interpretations. There are potential

costs or undesirable consequences associated with such a proposal.

First, in addition to all the other listed frames for numeration, morphological adjustments and Vocabulary Insertion, we now have to list for every item a subcategorization frame, determining the allosemantic variants of the functional items in the context of encyclopedic insertion. This adds an additional layer of complexity at the level of the semantic spell-out.

Second, in this approach, the mapping between syntactic structure and meaning can no longer be relied upon. In a framework like the exo-skeletal model (Borer, 2005), polysemy is possible in the domain of *words*, but not in functional structures; each structure has defined properties, imposing relatively strict conditions on its interpretation. An allosemantic approach, on the other hand, reduces the chance of both semantic and syntactic bootstrapping for the child acquiring their language.

Despite these potential pitfalls, namely the greater complexity of semantic listings associated to each root or functional head and its inevitable effect on learning, this approach has considerable advantages. First, by allowing a certain degree of autonomy to the syntax while establishing strict conditions for certain meanings to emerge in specific contexts, we are able to account for systems like that of Arabic number, where the same feature fills one function in a paradigm, and another function in a parallel one. For instance, instead of analyzing the feminine morpheme as polysemous, we can assume an allosemantic effect, where each meaning is determined by a specific syntactic context. In fact, by approaching the variable meaning of the feminine as a matter of polysemy, we rely on random associations between a certain feature and the multiple meanings associated to it, thus overlooking a crucial general-

ization regarding the role of the local syntactic context in assigning an interpretation to abstract morphemes. An allosemantic approach demonstrates the speakers' ability to acquire rule-governed behaviour, by presupposing a linguistic ability which is fundamentally computational, with as small as possible an inventory of idiosyncratic information appended on it, beyond the arbitrary pairing of sound and meaning. Applied to the data at hand, the allosemantic approach provides an elegant account of an intricate number system involving the coexistence of two paradigms, each one using the [+FEM] feature for a distinct and specific purpose. As I show in [Section 5.4](#), with this system, a child only needs to learn the grammatical category of each noun, and the interpretation of the feminine and plural morphemes results from the knowledge of this category.

The bottom line is that allosemy does make syntactic structure and meaning less reliable in addition to complicating the list of possible outcomes at LF. However, with a complex number system like the one we find in Arabic, and given the diachronic traces of the different functions of the feminine that are still found in the contemporary dialects, these drawbacks are a low price to pay in exchange for a model that allows us to restrict the associations between abstract morphemes and their multiple meanings by establishing well-defined syntactic contexts based on noun classification.

Having established a rationale and a proper definition for allosemy, I now turn to instances of allosemantic morphemes within the DP domain, which is the focus of the present thesis.

5.3 Allosemes in the DP

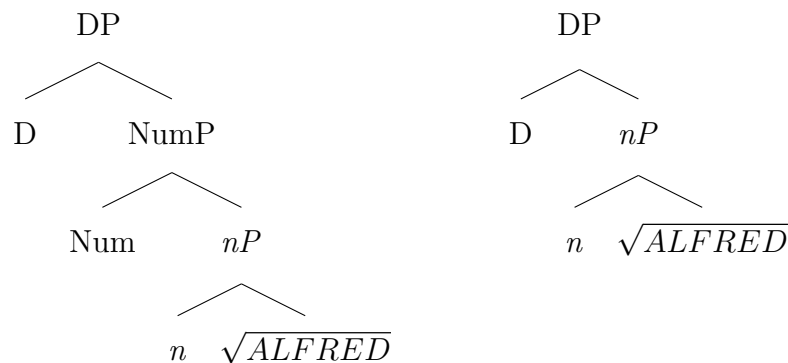
5.3.1 Saab and Lo Guercio (2019): Names as allosemes

Saab and Lo Guercio (2019) assume an allosemic relation between the different possible interpretations of names to account for data such as (6), where proper names, i.e., *Alfred* and *Orson* occur as predicates.

- (6)
- a. Alfreds are usually good filmmakers.
 - b. The Orson who filmed *A touch of Evil* is a good filmmaker.
 - c. The studio hired an Orson and an Alfred.
 - d. Every Alfred in the studio is a valuable employee.
 - e. There are two Alfreds in the studio.

The main claim is that whether a proper name is a predicate or has referential use depends on its immediate syntactic context. The referential use is the result of the proper name being selected by D, while the predicate interpretation results from a more complex structure including a Num projection, as illustrated in (7).

- (7) a. The syntax of predicate uses: b. The syntax of referential uses:



This analysis implies, following DM, that proper nouns do not come from the lexicon specified as referential, but are instead first categorized as nouns upon combination of the root with n . Then at LF, the complex nP is realized either as an e,t object (7-a), or as an entity-denoting object, e (7-b). Therefore, assuming the syntax in (7) for predicative and referential occurrences of a give name root, we end up with the following semantic realizations for the root *Alfred*:

- (8) LF instructions: semantic realizations of $\sqrt{\text{ALFRED}}$
- a. $[n+\sqrt{\text{ALFRED}}] \leftrightarrow$ “being an alfred”/Num_____
 - b. $[n+\sqrt{\text{ALFRED}}] \leftrightarrow$ “being called ALFRED”/D_____

The implicit idea behind this proposal is that referential DPs of the appropriate type encode referentiality in different ways depending on the absence or presence of Num. When Num is present, the structure below D denotes a predicate, so referentiality is encoded directly on D. In cases where Num is absent, the nP domain bears the referentiality information and D is expletivized. This captures in a different way Longobardi’s (1994) original proposal (see also Borer 2005).

The absence of Num, in turn, is justified by the absolute impossibility of pluralizing bare referential names (9).

- (9) *Alfreds went to the party.

To sum up, the proposal by Saab and Lo Guercio (2019) shows an example of how allosemy can account for the different semantic realizations of an nP based on its

neighbouring syntactic context.

5.3.2 The Num projection in Dagaare

Another example of allosemy within the DP comes from the interpretation of the suffix *-ri* in Dagaare (Grimm, 2012), which is the phonological realization of the Num head. Taken in isolation, this morpheme does not have a fixed semantic denotation. The latter is determined by the nature of the base to which the suffix is attached. Consider the paradigm presented in (10). The data exhibits a near minimal pair where the nouns ‘child’ and ‘seed’ share the same stem, yet the content of Num, characterized by the suffix *-ri*, marks the plural interpretation for ‘child’ and the singular interpretation for ‘seed’.

	Gloss	Singular	Plural	Stem
(10)	‘Child’	<i>bíé</i>	<i>bíírí</i>	bi-
	‘Seed’	<i>bírí</i>	<i>bíè</i>	bi-

(Grimm, 2014)

This system for marking number in Dagaare is initially surprising. However, if we adopt an allosemic treatment of the content of Num by determining its meaning in function of the base, things suddenly become less arbitrary. This is the account that I pursue here.

Dagaare has a collective system, just like Arabic and many other languages, where the collective base is morphologically unmarked and individual denotation takes an affix. This system occurs, of course, in parallel with a regular “count” system where reference to sums is morphologically marked. There are hence two

classes of nouns, each one belonging to one of the systems presented above. Grimm (2012) argues that the organizing force behind the nominal system of Dagaare is based on semantic principles. Basically, there are nouns referring to entities that are distinct and countable², and nouns referring to entities that are non-distinct and uncountable. When *-ri* attaches to the former, we obtain a sums reading, and when the same affix attaches to the latter, the result is an individual (singulative) interpretation. Put another way, the first group of nouns belong to the count system, while the second one belongs to the collective class.

In Chapter 3 I proposed, following many others (Borer, 2005; Mathieu, 2009, 2012, 2013; Borer and Ouwayda, 2010; Fassi Fehri, 2012, 2018b,a), that the singulative, as a unit counter is hosted on Num, and is hence in complementary distribution with the classifying plural. This suggests that the content of the functional head Num is allosemantic, with its interpretation being conditioned by the nature of the noun class, which is marked on *n*. This allosemantic relation is even more obvious in Dagaare, since the plural and the singulative are marked with the same affix - this suggests that Num has only one possible phonological realization. To account for this data, I propose that the singulative and the plural are allosemes of the Num head in Dagaare, which is realized by the VI *-ri* at PF. The interpretation of Num, realized as *-ri* is context-dependent, as illustrated in (11).

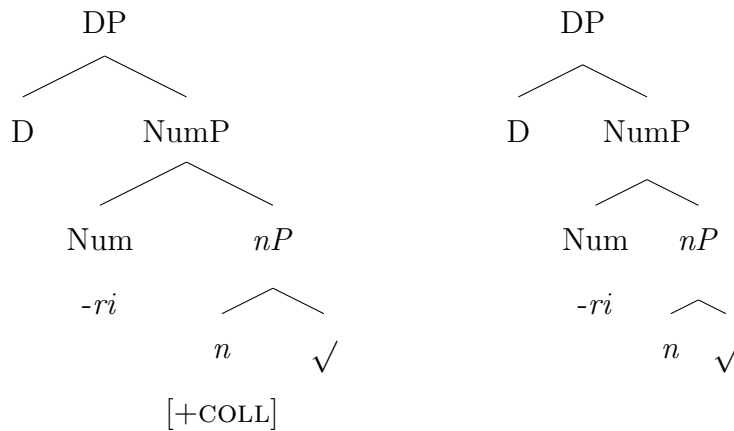
²Grimm establishes the notions of countability and individuation based on four factors: animacy, the likelihood of the referent of a noun to appear as a member of a pair or group (e.g. some body parts like kidneys are paired), distinguishability (e.g. it is easier to visually distinguish a bean than a grain of rice), and the canonical manner of interaction with a given entity (e.g. one can interact with berries one by one, but would scoop up grains of rice). The last two factors are based on Wierzbicka (1985) and were tested experimentally by Middleton et al. (2004).

(11) LF instructions: semantic realizations of Num

- a. [Num] ↔ “singulative” / _____ $n_{[+COLL]}$
 b. [Num] ↔ “plural” / _____ $n_{[-COLL]}$

As shown in (11), when attached to an nP marked [+COLL], the content of Num is interpreted as a singulative, as in (12-a). Otherwise, the content of Num is interpreted as a classifying plural, as illustrated in (12-b).

- (12) a. The singulative $-ri$: b. The plural $-ri$:



Having shown examples of how allosemy can be conditioned by functional heads within the DP, and more specifically how the nature of the features on n can condition the interpretation of the content of Num, I now turn to my proposal of an allosemic account of the multiple interpretations of certain functional projections within the TA noun. In this proposal, the nature of n not only conditions the semantic flavour of Num, but also the functional level of attachment of gender as well as its corresponding interpretation.

5.4 Proposal: Allosemantic morphemes in the TA DP

In this section, I put forward the proposal that the different instances of the feminine gender in TA are allosemantic. That is, the interpretation of the feminine depends on its syntactic context. Moreover, I extend my proposal to higher functional heads, claiming that the context-dependent interpretation of the feminine affects the semantics of different other nodes and abstract morphemes that are parasitic on the function of gender.

5.4.1 The feminine

The main puzzle pertaining to the different roles of gender features in TA is that their function depends on the class of the noun to which they are attached, that is, the presence or absence of the [+COLL] feature marked on *n*. Let us take the example of nouns with the [+FEM] feature. To recall the facts, consider the examples in (13), where feminine gender expresses noun categorization and female interpretation (13-a), noun categorization with no semantic contribution to sex interpretation (13-b), or individuation of a sums of entities expressed by a noun that is already syntactically formed (by masculine gender marking on *n* which marks the collective in Arabic).

- (13) a. fakrun-a [Tunisian Arabic]
turtle-FEM.SG
'a female turtle (as opposed to a male)'

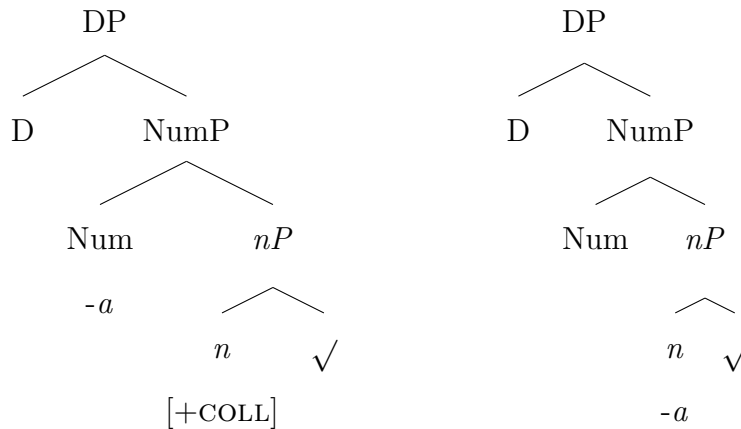
- b. tawla
table.FEM.SG
'a table'
- c. jormen-a
duck.-FEM.SING
'a duck (as opposed to a collection of ducks)'

We can observe in (13) that the interpretation of the [+FEM] feature is conditioned by the type of noun to which it attaches. Animate nouns of the count class constitute a syntactic context where the [+FEM] feature denotes a female entity (13-a). Inanimate nouns of the count class condition the [+FEM] feature to make it categorize the root as a noun without denoting any semantic gender in particular. Nouns of the collective class, whether animate or inanimate, constitute a syntactic context in which the [+FEM] feature denotes a single individual out of a collection. All the elements conditioning the interpretation of the [+FEM] feature are encoded on *n* and, to some extent, on the root. Whenever the [+FEM] feature is realized, the semantic component inspects the set of possible interpretations and retrieves the alloeme 'individual/singulative' whenever *n* is marked [+COLL]. Therefore, given that gender can be expressed on different heads in Arabic, as put forth in previous chapters, we can predict that when the *n* bears the [+COLL] feature, the [+FEM] gender is automatically interpreted on Num, as a marker of individuation. When *n* does not bear a [+COLL] feature, gender is interpreted on *n*, therefore it has a nominalizing function with a possible contribution to semantic gender interpretation, depending on whether the concept denoted by the root bears the concept of animacy. The conditioned interpretations of the [+FEM] feature are represented in (14).

- (14) LF instructions: semantic realizations of [+FEM]
- a. [+FEM] ↔ “singulative” / _____ $n_{[+COLL]}$
 - b. [+FEM] ↔ “nominalizer” / _____ $n_{INANIMATE}$
 - c. [+FEM] ↔ “nominalizer” and “female” / _____ $n_{ANIMATE}$

The corresponding structures are represented in (15).

- (15) a. The singulative feminine: b. The nominalizer feminine:



In sum, the feature values on n determine the place of attachment of gender and, by consequence, the interpretation thereof. This, I argue, is due to the fact that Arabic has two parallel number systems that make different uses of gender marking. In line with an allosemy analysis, the [+FEM] feature is an abstract one that is semantically underdetermined when considered in isolation. It is only in combination with an n valued with a class feature that gender takes a fixed semantic interpretation. Based on these facts, I propose that the singulative and semantic gender interpretations are allosemes of the abstract [+FEM] feature.

As a side note, it is important to clarify that the relevant criterion to identify an abstract node is the syntactic identity of the morpheme, not its phonological identity. For instance, I rely on agreement to identify the feminine morpheme in Arabic, and not on the fact that this morpheme is phonologically marked by the suffix *-a*.

Another important matter to address is the idea that there should be a principled link between the interpretations invoked by allosemy. For instance, a common semantic core is easily detected among the allosemes of proper names (Saab and Lo Guercio, 2019) and among those of the abstract root HOUSE (Harley, 2014). The same can be said of allomorphy, where it is often possible to reconstruct the logic behind a choice of allomorphs on the basis of phonological context (e.g., *-t/-d* alternations). In what follows, I explain the logic behind the less obvious interpretive link between the allosemes of the Arabic feminine.

While there is a common core in root allosemes, functional allosemes are motivated by a *functional* logic rather than an *interpretational* one. For instance, returning to the different allosemes associated with the *v* head in Icelandic (Wood, 2012), there is no obvious semantic connection between a stative and a dynamic interpretation. Instead, the link is found in the function associated with these allosemes: they turn the complement into a verb denoting a state or an event. For the case of the feminine in Arabic, we observe that this morpheme always marks a contrast with the base. This was also observed in the diachronic facts. For instance, if the base denotes a male entity, the feminine turns it into a female one. If the base expresses a collective, the feminine turns it into an individual. Therefore, although there is no clear common semantic core between a female-denoting feminine and a

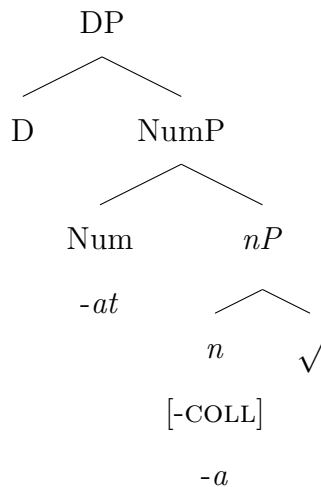
- (18) El fannan-a-at mtaʕ tunes el kol mawjud-een. [Tunisian Arabic]
 the artist-FEM-PL of Tunisia the all present-PL
 ‘All the female artists from Tunisia are present.’
- (19) #El toffeħ-a-at mtaʕ el marché central el kol tbeʕ-u.
 the apple-FEM.SING-PL of the marketplace central the all sold-3.PL
 ‘All the apples from the central marketplace were sold.’

The examples above show that when *-at* is suffixed to a singular count noun, the reading is a general plural (18), while when *-at* is attached to a singulative derived from a collective noun, we obtain a paucal reading (19). Both contexts in (18) and (19) require a plural that refers to more than a few entities. In (18), reference is made to all Tunisian female artists, a number that certainly exceeds ten. The use of the feminine sound plural is felicitous in this environment, since it refers to an unbounded sum of artists. In (19), reference is made to all apples of the central market, a number that also exceeds ten. In this case however, the use of the feminine sound plural is not felicitous, since *toffeħ* ‘apples’ is a collective noun in TA. Therefore, the feminine plural, realized as the suffix *-at* is attached to its singulative and denotes a paucal. In sum, the nature of the base noun conditions the interpretation and syntactic level of attachment of the feminine plural. The interpretation of *-at* is thus contextually allosemantic, as illustrated in (20).

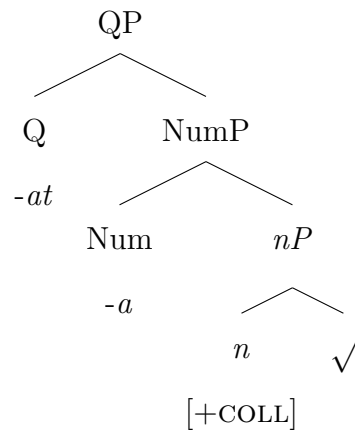
- (20) LF instructions: semantic realizations of [+FEM,+PL]
- a. [+FEM,+PL] ↔ “classifying plural” / _____ $n_{[+FEM]}$
- b. [+FEM,+PL] ↔ “paucal” / _____ $[_{Num-a}]$

As shown in (20), when attached to an *nP* marked [+FEM], the [+FEM, +PL] feature is interpreted as a classifying plural on Num, as in (21-a). When attached to a NumP marked [+FEM], the [+FEM, +PL] feature marks a paucal interpretation and is hosted on Q, as illustrated in (21-b).

(21) a. The plural:



b. The paucal:



In the context of nouns that are not marked [+COLL] on *n*, the feminine plural is interpreted as an inclusive plural/a classifier and is hosted on Num. The same feature, when merged to an *nP* with a [+COLL] feature on *n* is interpreted as a paucal and, as such, is hosted on the Q head, since it is, in my view (see also [Mathieu 2013, 2014](#); [Dali and Mathieu 2019a](#)) an operation on a singulative base where gender is hosted and interpreted on Num. Essentially, we can observe that the interpretation of the [+FEM, +PL] feature bundle depends primarily on the place of attachment and interpretation of the [+FEM] feature. This corroborates the idea that the role of the feminine gender is multifunctional in Arabic and affects other levels of the DP,

such as number.

5.4.3 Summary

To summarize, I appealed to the notion of allosemy in order to account for a system where the semantic interpretations of the feminine and plural affixes depend on the syntactic contexts in which they appear. More specifically, I argued that the interpretation of the feminine morpheme, phonologically realized as the affix *-a* depends on whether the morpheme attaches to a noun marked [+COLL] on *n* or not. In the former case, the feminine is a singulative and is hosted on Num. In the latter, the feminine is a nominalizer and is on *n* (where it can be interpretable or uninterpretable depending on the animate/inanimate nature of the referent). Consequently, the interpretation of the feminine plural marker depends on the place of attachment of the feminine. When attached to *-a* hosted in *n*, *-at* is interpreted as an inclusive plural and is on Num. However, when *-at* is affixed to an *-a* that is on Num (i.e., a singulative), it is a paucal that is hosted on Q. Put another way, the interpretation of *-at* is also dependent on the nature of the noun (count or collective), since the latter determines the interpretation of the feminine gender that affects the shape of the plural (*-at* instead of *-een*). This analysis accounts for a system with two parallel number paradigms, causing class marking to condition the interpretations of gender and number morphemes.

5.5 Conclusion

In this chapter, I looked at the notion of contextual allosemy, which I defined as the phenomenon in which a single morpheme can have multiple semantic realizations - the equivalent of allomorphy in the LF side of the spell-out (following [Wood 2012](#); [Marantz 2013](#); [Harbour 2014](#); [Wood and Marantz 2017](#)). I established that allosemy was the simplest way to account for the effects of an intricate number system with inverse paradigms on the interpretation of a gender morpheme that plays different roles depending on the category of the noun.

In addition to existing analyses in the domain of roots and functional heads (e.g., *v*), I showed that an allosemic account was also tenable within the DP by presenting two cases. The first example was the treatment of proper names by [Saab and Lo Guercio \(2019\)](#). The main claim is that whether a proper name is a predicate or has a referential use is conditioned by its immediate syntactic context: when the *nP* is selected by D, it gets a referential interpretation, and when the same *nP* is part of a more complex structure including a Num projection, we obtain a predicate interpretation. The second example is the interpretation of the content of Num (phonologically realized as *-ri*) in Dagaare (based on an analyses by [Grimm 2012](#)). When attached to a collective *nP*, the content of Num is interpreted as a singulative operator, and when it is attached to a count *nP*, the content of Num is interpreted as a plural.

Based on these previous analyses suggesting that the meaning of a morpheme in the DP domain can be conditioned by the structural context of the morpheme, I proposed that the different interpretations associated with the [+FEM] feature in TA

are best described as allosemantic in nature. In sum, the feminine *-a* attached to a noun not marked [+COLL] on *n* has the function of a nominalizer, and whether the feminine feature is semantically interpretable or not depends on the animacy of the entity of reference. On the other hand, the feminine morpheme attached to a noun marked [+COLL] on *n* is interpreted on Num and translates as a singulative. Furthermore, the interpretation and level of attachment of the [+FEM] feature defines the context of interpretation of the feminine plural affix *-at*. When attached to a *nominalizer* [+FEM], *-at* is interpreted as a classifying plural and is hosted on Num. However, when attached to a *singulative* [+FEM], *-at* is a paucal and is hosted on Q. The model proposed in this chapter shows how the classification of a noun (count versus collective) is at the root of the function of gender and number markers in TA. This emphasizes the importance of defining the category of a noun as a formal feature on a functional head like *n*, which allows us to manipulate the DP structure and its different dependencies in a restricted way, allowing us to obtain the appropriate meaning.

Chapter 6

Conclusion

The present thesis has argued for a non-unified morpho-syntactic analysis of gender and number morphemes in TA. Gender and number markers in TA can take different interpretations depending on the base noun they attach to. This is due to the fact that TA has two parallel noun categories (count and collective), each one using gender features in a different way. Based on the co-existence of these two paradigms, a unified treatment of gender and number would not account for the conditioned interpretations of gender and number markers in TA.

To account for the dependency between noun category and gender/number marker interpretation, I used the concept of contextual allosemy, whereby a single morpheme can have multiple semantic realizations, based on its morphological context. The basic claim is that when attached to a count noun, gender has a nominalizing function and is hosted on *n*. When attached to a collective noun, gender (most specifically, the feminine) has a dividing function and is hosted on Num. Similarly, the interpre-

tation of the feminine plural marker *-at* depends on whether it is attached to a count or a collective noun. When attached to count noun, *-at* is interpreted as a plural, whereas when it is attached to a collective noun, it has the a paucal interpretation.

A thorough analysis of the collective class was conducted, based on the distribution and semantic properties of collective nouns. It was concluded that the collective class is not a semantic but rather a morphological class. Therefore, the class is specified on *n*.

Another claim that was made in this thesis is that broken plurals are inflectional, based on the fact that they are productive and do not give rise to special meaning.

The interpretation of gender morphemes in other languages with a singulative system will be investigated in future research, using the notion of allosemy.

Bibliography

- Abd-Rabbo, M. Z. (1990). Sound plural and broken plural assignment in classical arabic. *Perspectives on Arabic linguistics I*, 55–94.
- Abdel-Massih, E. T., Z. N. Abdel-Malek, and E.-S. M. Badawi (1981). *A reference grammar of Egyptian Arabic*. Center for Near Eastern and North African Studies, Univ. of Michigan.
- Acquaviva, P. (2008). *Lexical plurals: A morphosemantic approach*. Oxford: Oxford University Press.
- Acquaviva, P. (2009). Roots and lexicality in distributed morphology. In *Fifth York-Essex Morphology Meeting (YEMM), 9th February and 10th February 2008, Department of Language and Linguistic Science, University of York*. University of York. Department of Language and Linguistic Science.
- Acquaviva, P. (2015). Singulatives. *Word-Formation: An International Handbook of the Languages of Europe. HSK Handbücher zur Sprach-und Kommunikationswissenschaft*. Berlin: de Gruyter Mouton, 1171–82.

- Aikhenvald, A. Y. (2000). *Classifiers: A typology of noun categorization devices: A typology of noun categorization devices*. OUP Oxford.
- Aikhenvald, A. Y. (2003). Classifiers in spoken and in signed languages: how to know more. In *Perspectives on classifier constructions in sign languages*, pp. 97–100. Psychology Press.
- Aksenov, A. (1984). On the problem of the extralinguistic motivation of grammatical gender. *Voprosy jazykoznanija* 1, 14–25.
- Alexiadou, A. (2004). Inflection class, gender and dp internal structure. *Explorations in nominal inflection*, 21–50.
- Arad, M. (2003). Locality constraints on the interpretation of roots: The case of hebrew denominal verbs. *Natural Language & Linguistic Theory* 21(4), 737–778.
- Arad, M. (2005). *Roots and patterns: Hebrew morpho-syntax*, Volume 63. Springer Science & Business Media.
- Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*. Number 22. MIT press.
- Bale, A., M. Gagnon, and H. Khanjian (2011). On the relationship between morphological and semantic markedness: The case of plural morphology. *Morphology* 21, 197–221.
- Bale, A. C. and D. Barner (2009). The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26(3), 217–252.

- Barker, C. (1992). Group terms in English: Representing groups as atoms. *Journal of Semantics* 9, 69–93.
- Belnap, K. (1991). *Grammatical agreement variation in Cairene Arabic*. Ph. D. thesis, University of Pennsylvania.
- Bennett, M. (1974). *Some Extensions of a Montague Fragment of English*, University of California at Los Angeles. Ph. D. thesis, PhD. dissertation.
- Bernstein, J. (1993). *Topics in the syntax of nominal structure across Romance*. Ph. D. thesis, CUNY.
- Blench, R. (2013). Introduction to the teme in languages. In *Nuba Mountain Language Studies*. Rudiger Koppe.
- Bobaljik, J. D. (2000). The ins and outs of contextual allomorphy. *University of Maryland working papers in linguistics* 10, 35–71.
- Bonet, E. (1991). *Morphology after syntax: Pronominal clitics in Romance*. MIT.
- Borer, H. (2005). *Structuring sense: Volume 1: In name only*. Oxford: Oxford University Press.
- Borer, H. and S. Ouwayda (2010). Men and their apples: Dividing plural and agreement plural. Paper presented at GLOW in Asia VIII, Beijing Language and Culture University.
- Boudelaa, S. and M. G. Gaskell (2002). A re-examination of the default system for arabic plurals. *Language and cognitive processes* 17(3), 321–343.

- Brustad, K. (2000). *The syntax of spoken Arabic*. Georgetown University Press.
- Butler, L. (2012). Crosslinguistic and experimental evidence for non-number plural. *Linguistic Variation 12*, 27–56.
- Carlson, G. N. (1977). A unified analysis of the english bare plural. *Linguistics and philosophy 1*(3), 413–457.
- Cheng, C.-Y. (1973). Response to moravcsik.
- Cheng, L. L.-S. and R. Sybesma (1999). Bare and not-so-bare nouns and the structure of np. *Linguistic inquiry 30*(4), 509–542.
- Chierchia, G. (1998). Plurality of mass nouns and the notion of “semantic parameter”. In *Events and grammar*, pp. 53–103. Springer.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Ma: MIT Press.
- Chomsky, N. (2000). Minimalist inquiries: The framework (mitopl 15). *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155.
- Chomsky, N. (2001). Derivation by phase. ken hale: A life in language, ed. by michael kenstowicz, 1-52.
- Cohen, M. (1970). *Traité de langue amharique*.
- Comrie, B. (1999). Grammatical gender systems: a linguist’s assessment. *Journal of Psycholinguistic research 28*(5), 457–466.
- Corbett, G. (2000). *Number*. Cambridge University Press.

- Corbett, G. (2011). Systems of gender assignment. In M. S. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online. Max Planck Digital Library, Munich.*, pp. <http://wals.info/chapter/32>.
- Corbett, G. (2015). Hybrid nouns and their complexity. In J. Fleischer, E. Rieken, and P. Widmer (Eds.), *Agreement from a diachronic perspective*, pp. 191–214. Berlin: Mouton de Gruyter.
- Corbett, G. G. (1991). *Gender*. Cambridge Textbooks in Linguistics. Cambridge University Press.
- Corbett, G. G. (1996). Minor number and the plurality split: 397. *Italian journal of linguistics* 8(1), 101–122.
- Corbett, G. G. (2006). Gender, grammatical. *The Encyclopedia of language and linguistics. 2nd Edition. Oxford: Elsevier*, 749–756.
- Cowell, M. (1964). *A reference grammar of Syrian Arabic (based on the dialect of Damascus)*. Washington, D.C.: Georgetown University Press.
- Croitor, B. and I. Giurgea (2009). On the so-called romanian neuter. *Bucharest Working Papers in Linguistics* 11(2), 21–39.
- Dahl, O. et al. (2000). Animacy and the notion of semantic gender. *Trends in linguistics studies and monographs* 124, 99–116.
- Dali, M. (2015). The feminine operator in arabic. Master's thesis, University of Ottawa.

- Dali, M. and E. Mathieu (2016). Les pluriels internes féminins de l'arabe tunisien. *Linguisticae Investigationes* 59, 253–271.
- Dali, M. and E. Mathieu (2019a). Broken plurals and (mis)matching of ϕ -features in tunisian arabic In review.
- Dali, M. and E. Mathieu (2019b). Historical change in sub-word formation: The case of Arabic -a(t). *In prep.*.
- Dali, M. and E. Mathieu (2020). Broken plurals and (mis)matching of ϕ -features in tunisian arabic.
- Danon, G. (2011). Agreement and DP-internal feature distribution. *Syntax* 14, 297–317.
- Danon, G. (2013). Agreement alternations with quantified nominals in Modern Hebrew. *Journal of Linguistics* 49, 55–92.
- De Swart, H. and J. Zwart (2009). Less form: more meaning. *Lingua* 119 2, 280–295.
- Deal, A. R. (2017). Countability distinctions and semantic variation. *Natural Language Semantics* 25(2), 125–171.
- den Dikken, M. (2001). Plurilinguals, pronouns and quirky agreement. *The Linguistic Review* 18, 19–41.
- Dimmendaal, G. (1983). *The Turkana language*. Dordrecht: Foris Publications.
- Dimmendaal, G. J. (2000). Number marking and noun categorization in nilo-saharan languages. *Anthropological linguistics*, 214–261.

- Dixon, R. M. W. (1982). *Where have All the Adjectives Gone? And Other Essays in Semantics and Syntax*. Berlin, De Gruyter Mouton.
- Dixon, R. M. W. (1986). Noun classes and noun classification in typological perspective. *Noun classes and categorization*, 105–112.
- Dressler, W. U. (1998). Produktivität als zentralbegriff der gramatischen (bes. flexionsmorphologischen) systemadäquatheit. terzan-kopeccky, l. *Sammel-band des II. Internationalen Symposions zur Natürlichkeitstheorie*, 113–124.
- Embick, D. and R. Noyer (1999). Locality in post-syntactic operations. *MIT working papers in linguistics* 34(265-317).
- Embick, D. and R. Noyer (2007). Distributed morphology and the syntax/morphology interface. *The Oxford handbook of linguistic interfaces* 289324.
- Erwin, W. M. (2004). *A short reference grammar of Iraqi Arabic*. Georgetown University Press.
- Farkas, D. F. (1990). Two cases of underspecification in morphology. *Linguistic Inquiry* 21(4), 539–550.
- Fassi Fehri, A. (1984). Agreement, binding and coherence. paper given at the conference on agreement in natural language.
- Fassi Fehri, A. (1988). Agreement in Arabic, binding and coherence. In M. Barlow and C. Ferguson (Eds.), *In Agreement in Natural Language: Approaches, Theories, Description*, pp. 107–158.

- Fassi Fehri, A. (2004). Nominal classes and parameters across interfaces and levels, with a particular reference to Arabic. *Linguistic Variation Yearbook 4*, 41–108.
- Fassi Fehri, A. (2012). *Key features and parameters in Arabic grammar*. Amsterdam: John Benjamins.
- Fassi Fehri, A. (2018a). *Constructing Feminine to Mean: Gender, Number, Numeral, and Quantifier Extensions in Arabic*. Rowman & Littlefield.
- Fassi Fehri, A. (2018b). Multiple facets of constructional arabic gender and ‘functional universalism’ in the dp.
- Ferguson, C. A. (1989). Grammatical agreement in Classical Arabic and the modern dialects: A response to versteegh’s pidginization hypothesis. *Al-’Arabiyya 22 (1–2)*, 5–18.
- Ferrari, F. (2005). *A syntactic analysis of the nominal systems of Italian and Luganda: How nouns can be formed in the syntax*. New York University.
- Fleisch, H. (1961). *Traité de philologie arabe*. vol. 1 préliminaires, phonétique morphologie nominale.
- Galani, A. (2004). So, word formation: Syntax or morphology? formation of non-suppletive stems in modern greek’. In *Proceedings of the III International Conference of Greek Linguistics*.
- Gaudefroy-Demombynes, M. and R. Blachère (1952). *Grammaire de l’arabe classique. Troisième Édition*. Paris: G.-P. Maisonneuve Larose.

- Giurgea, I. (2008). Gender on definite pronouns. *Bucharest Working Papers in Linguistics* (1), 97–121.
- Goddard, I. (2002). Grammatical gender in Algonquian. In H. C. Wolfart (Ed.), *Papers of the thirty-third Algonquian conference*, pp. 185–231.
- Greenberg, J. (1966). *Universals of language*. MIT Press.
- Greenberg, J. H. (1972). Numeral classifiers and substantive number: problems in the genesis of a linguistic type. *Working Papers on Language Universals* 9, 2–39.
- Greenberg, J. H. (1978). How does a language acquire gender markers. *Universals of human language* 3, 47–82.
- Grice, H. P. (1967). Logic and conversation. William James lectures, Harvard University. *Studies in syntax* 3.
- Grimm, S. (2012). Inverse number marking and individuation in Dagaare. In D. Masmam (Ed.), *Count and mass across languages*, pp. 75–98. Oxford: Oxford University Press.
- Grinevald, C. (2000). A morphosyntactic typology of classifiers. *Systems of nominal classification* 4, 50–92.
- Haelewyck, J.-C. (2016). *Grammaire comparée des langues sémitiques*. Bruxelles, éditions Safran.
- Halle, M. (1997). Fission and impoverishment. *MIT Working Papers in Linguistics* 30, 425–449.

- Halle, M. and A. Marantz (1993). Distributed morphology and the pieces of inflection. hale, k. & sj keyser (eds.), *the view from building 20*.
- Halle, M. and A. Marantz (2004). Distributed morphology and. *Morphology: Critical Concepts in Linguistics 1*, 381.
- Hammerly, C. (2018). Limiting gender. In M. D. Éric Mathieu and G. Zareikar (Eds.), *Gender and noun classification*, pp. 93–118. Oxford: Oxford University Press.
- Hammond, M. (1988). Templatic transfer in arabic broken plurals. *Natural Language & Linguistic Theory 6*(2), 247–270.
- Harbour, D. (2008). Mass, non-singularity, and augmentation. *MIT Working papers in linguistics 49*, 239–266.
- Harbour, D. (2014). Paucity, abundance, and the theory of number. *Language 90*, 185–229.
- Harley, H. (2008). When is a syncretism more than a syncretism? impoverishment, metasyncretism, and underspecification. *Phi theory: Phi-features across modules and interfaces*, 251–294.
- Harley, H. (2014). On the identity of roots. *Theoretical linguistics 40*(3-4), 225–276.
- Harley, H. and R. Noyer (1998). Licensing in the non-lexicalist lexicon: nominalizations, vocabulary items and the encyclopedia. *MIT working papers in linguistics 32*, 119–137.

- Harley, H. and R. Noyer (1999). Distributed morphology. *Glott international* 4(4), 3–9.
- Harley, H. and R. Noyer (2000). Licensing in the non-lexicalist lexicon. *The lexicon/encyclopaedia interface*, 349–374.
- Harris, J. W. (1991). The exponence of gender in spanish. *Linguistic Inquiry* 22(1), 27–62.
- Hasselbach, R. (2014a). Agreement and the development of gender in Semitic (Part I). *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 164, 33–64.
- Hasselbach, R. (2014b). Agreement and the development of gender in Semitic (Part II). *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 164, 319–344.
- Hieda, O. (2006). On singulative languages in Nilotic languages. *Nilo-Ethiopian Studies* 10, 1–14.
- Hockett, C. F. (1958). *A Course in Modern Linguistics*. Macmillan, New York.
- Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: q-based and r-based implicature. In D. Schiffrin (Ed.), *Meaning, Form, and Use in Context: Linguistic Applications*, pp. 11–42. Washington, D.C.: Georgetown University Press.
- Keenan, C. (2013). ‘a pleasant three days in philadelphia’: Arguments for a pseudopartitive analysis. *University of Pennsylvania Working Papers in Linguistics* 19(1), 11.

- Kihm, A. (2003). Les pluriels internes de l'arabe: système et conséquences pour l'architecture de la grammaire. *Recherches linguistiques de Vincennes* (32), 109–156.
- Kihm, A. (2005). Noun class, gender, and the lexicon-syntax-morphology interfaces: A comparative study of Niger-Congo and Romance languages. In G. Cinque and R. S. Kayne (Eds.), *The Oxford handbook of comparative syntax*, pp. 459–512. Oxford University Press.
- King, G. (2015). *Modern Welsh: a comprehensive grammar*. Routledge.
- Kramer, R. (2009). *Definite markers, phi-features, and agreement: A morphosyntactic investigation of the Amharic DP*. Ph. D. thesis, UC Santa Cruz.
- Kramer, R. (2014). Gender in Amharic: A morphosyntactic approach to natural and grammatical gender. *Language Sciences* 43, 202–115.
- Kramer, R. (2015). *The Morphosyntax of Gender: Evidence from Amharic*. Oxford University Press.
- Kramer, R. (2016). A split analysis of plurality: Number in amharic. *Linguistic Inquiry* 47(3), 527–559.
- Kramer, R. and L. Winchester (2018). Number and gender agreement in saudi arabic: Morphology vs. syntax1. *Mayowa Akinlotan*, 39.
- Krifka, M. (1989). Nominal reference, temporal constitution and quantification. In R. Bartsch, J. van Benthem, and P. van Emde Boas (Eds.), *Semantics and contextual expression*, pp. 75–111. Dordrecht: Springer.

- Lahrouchi, M. and N. Lampitelli (2014). On plurals, noun phrase and num (ber) in moroccan arabic and djibouti somali. *The Form of Structure, the Structure of Form: Essays in honor of Jean Lowenstamm* 12.
- Lahrouchi, M. and R. Ridouane (2016). On diminutives and plurals in moroccan arabic. *Morphology*, 1–23.
- Landau, I. (2015). Dp-internal semantic agreement: A configurational analysis. *Natural Language and Linguistic Theory* 34, 975–1020.
- Landman, F. (1989). Groups i and ii. *Linguistics and Philosophy* 12, 559–605 and 723–744.
- Lecarme, J. (2002). Gender ‘polarity’: Theoretical aspects of somali nominal morphology. *Many morphologies* 109, 141.
- Leslau, W. (1995). *Reference grammar of Amharic*. Otto Harrassowitz Verlag.
- Levy, M. M. (1971). *The plural of the noun in Modern Standard Arabic*. Ph. D. thesis, uitgever niet vastgesteld.
- Link, G. (2002). The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Formal semantics: The essential readings*, 127–146.
- Lipiński, E. (2001). *Semitic languages outline of a comparative grammar. Second edition*. Leuven, Peeters.
- Longobardi, G. (1994). Reference and proper names: A theory of n-movement in syntax and logical form. *Linguistic inquiry*, 609–665.

- Lowenstamm, J. (2008). Little n, root, and types of nouns. *The Sounds of Silence: Empty Elements in Syntax and Phonology*, 105–143.
- Lowenstamm, J. (2012). Feminine and gender, or why the ‘feminine’ profile of french nouns has nothing to do with gender. *Linguistic Inspirations. Edmund Gussmann in memoriam, Eugeniusz Cyran, Henryk Kardela, and Bogdan Szymanek (eds.)*, Wydawnictwo Katolicki Uniwersytet Lubelski, Lublin, 371–406.
- Marantz, A. (1997). No escape from syntax: Don’t try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania working papers in linguistics* 4(2), 14.
- Marantz, A. (2001). Words. *WCCFL XX Handout, USC*.
- Marantz, A. (2009). Roots, re-, and affected agents: can roots pull the agent under little v. In *workshop Roots: Word Formation from the Perspective of “Core Lexical Elements”*, Stuttgart.
- Marantz, A. (2013). Verbal argument structure: Events and participants. *Lingua* 130, 152–168.
- Matasović, R. (2004). *Gender in Indo-European*. Universitaetsverlag Winter.
- Mathieu, E. (2009). On the mass/count distinction in Ojibwe. Article presented at the Mass/count workshop, organized by Diane Massam, University of Toronto. Feb. 7-8.
- Mathieu, E. (2012). Flavors of division. *Linguistic Inquiry* 43, 650–679.

- Mathieu, E. (2013). On the plural of the singulative. *McGill Working Papers in Linguistics 23*.
- Mathieu, E. (2014). Many a plural. In A. Aguilar-Guevara, B. L. Bruyn, and J. Zwarts (Eds.), *Weak referentiality*, pp. 157–181.
- Matushansky, O. (2013). Gender confusion.
- McCarthy, J. and A. Prince (1990a). Foot and word in prosodic morphology: The arabic broken plural. *Natural Language And Linguistic Theory 8 (2)*, 209–283.
- McCarthy, J. and A. Prince (1990b). Prosodic morphology and templatic morphology. In *Perspectives on Arabic linguistics II: papers from the second annual symposium on Arabic linguistics*, pp. 1–54. John Benjamins Amsterdam.
- McCawley, J. D. (1975). Lexicography and the count-mass distinction. In *Annual Meeting of the Berkeley Linguistics Society*, Volume 1, pp. 314–321.
- Morin, Y. C. (1986). A morphological convergence between consonant liaison and schwa deletion in the picard and walloon dialects of french. *Andersen (1986: 211-222)*.
- Mous, M. et al. (2008). Number as an exponent of gender in cushitic. *Typological Studies in Language 75*, 137.
- Murtonen, A. (1964). *Broken plurals, origin and development of the system*. Brill Archive.

- Nevins, A. (2011). Phonologically conditioned allomorph selection. *The Blackwell companion to phonology*, 1–26.
- Nichols, J. (1992). *Linguistic diversity in space and time*. University of Chicago Press.
- Nougayrol, P. (1990). La langue des aiki dits rounga: Tchad, république centrafricaine: esquisse descriptive et lexicque.
- Noyer, R. (1998). Impoverishment theory and morphosyntactic markedness. morphology and its relation to phonology and syntax.
- Ojeda, A. E. (1992). The semantics of number in arabic. In *Semantics and Linguistic Theory*, Volume 2, pp. 303–326.
- Ouwayda, S. (2014). *Where Number Lies: Plural marking, numerals, and the collective-distributive distinction*. University of Southern California.
- Pearson, H. (2011). A new semantics for group nouns. In *Proceedings of WCCFL*, Volume 28, pp. 160–168.
- Pelletier, R. (1989). *Gender Agreement in Telugu, Kurukh, and Malto*. MA dissertation, Yale University, New Have, Connecticut.
- Pesetsky, D. and E. Torrego (2004). *Tense, case, and the nature of syntactic categories*. na.
- Picallo, M. C. (1991). Nominals and nominalizations in catalan. *Probus* 3(3), 279–316.

- Picallo, M. C. (2008). Gender and number in romance. *Lingue e linguaggio* 7(1), 47–66.
- Quine, W. V. O. (1960). Word and object: An inquiry into the linguistic mechanisms of objective reference.
- Ratcliffe, R. R. (1998). *The broken plural problem in Arabic and comparative Semitic: allomorphy and analogy in non-concatenative morphology*, Volume 168. John Benjamins Publishing.
- Ritter, E. (1991). Two functional categories in noun phrases: Evidence from modern hebrew. In *Perspectives on phrase structure: Heads and licensing*, pp. 37–62. Brill.
- Ritter, E. (1993). Where is gender? *Linguistic Inquiry*, 795–803.
- Rothstein, S. (2010). Counting and the mass/count distinction. *Journal of semantics* 27(3), 343–397.
- Saab, A. and N. Lo Guercio (2019). No name: The allosemy view. *Studia Linguistica*.
- Sauerland, U. (2003). A new semantics for number. In R. Youn and Y. Zhou (Eds.), *Proceedings of semantics and linguistic theory (SALT 13)*, pp. 258–275. Ithaca, NY: CLC publications.
- Sauerland, U., J. Andersen, and K. Yatsuhira (2005). The plural is semantically unmarked. In S. Kepser and M. Reis (Eds.), *Linguistic evidence: Empirical, theoretical and computational perspectives*, pp. 413–433. Berlin: Mouton de Gruyter.

- Schwarzschild, R. (2011). Stubborn distributivity, multiparticipant nouns and the count/mass distinction. In *Proceedings of NELS*, Volume 39, pp. 661–678. Graduate Linguistics Students Association, University of Massachusetts
- Siddiqi, D. (2009). *Syntax within the word: Economy, allomorphy, and argument selection in Distributed Morphology*, Volume 138. John Benjamins Publishing.
- Smith, P. (2015). *Feature mismatches: Consequences for syntax, morphology and semantics*. Doctoral dissertation, University of Connecticut.
- Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher order implicature. In U. Sauerland and P. Stateva (Eds.), *Presupposition and implicature in compositional semantics*, pp. 243–281. Basingstoke: Macmillan.
- Steriopolo, O. and M. Wiltschko (2010). *Distributed GENDER hypothesis*. na.
- Stickney, H. (2009). The emergence of dp in the partitive structure.
- Stolz, T. (2001). Singulative-collective: Natural morphology and stable cases in welsh number inflexion of nouns. *Sprachtypologie und Universalienforschung* 54(1), 52–76.
- Talmoudi, F. (1980). The arabic dialect of szisa. *Acta Universitatis Gothoburgensis: Orientalia Gothoburgensia* 4.
- Treis, Y. (2014). Number in kambaata.
- Trépos, P. (1956). *Le pluriel breton*. Rennes: Imprimeries Réunies.

- Velupillai, V. (2012). *An introduction to linguistic typology*. John Benjamins Publishing.
- Wechsler, S. and L. Zlatić (2003). *The many faces of agreement*. Stanford: CLSI.
- Weiss, M. (1845). *A Short Grammar of the Maltese Language*. M. Weiss.
- Wiltschko, M. (2008). The syntax of non-inflectional plural marking. *Natural Language & Linguistic Theory* 26(3), 639–694.
- Wiltschko, M. (2012). Decomposing the mass/count distinction: Evidence from languages that lack it. *Count and mass across languages*, 120–146.
- Wood, J. (2012). *Icelandic morphosyntax and argument structure*. New York University New York, NY.
- Wood, J. and A. Marantz (2017). The interpretation of external arguments. *The verbal domain*, 255–278.
- Wright, W. (1933). *A Grammar of the Arabic Language (Volume I)*. Cambridge, MA: MIT Press.
- Wurzel, W. (1984). Flexionmorphologie und Natürlichkeit. (flexion-morphology and naturalness.) berlin.
- Wurzel, W. U. (1987). System-dependent morphological naturalness in inflection. *Leitmotifs in natural morphology* 10, 59–95.
- Zabbal, Y. (2002). *The semantics of number in the Arabic noun phrase*. MA dissertation, University of Calgary.

Zweig, E. (2009). Number-neutral bare plurals and the multiplicity implicature.

Linguistics and Philosophy 32, 353–407.