BRILL

Broken plurals and (mis)matching of ϕ-features in Tunisian Arabic

Myriam Dali
University of Ottawa, Ottawa, ON, Canada
mdali021@uOttawa.ca

Eric Mathieu
University of Ottawa, Ottawa, ON, Canada
emathieu@uOttawa.ca

Abstract

The aim of this paper is to explain an unusual agreement pattern that arises between Tunisian Arabic broken plurals and their targets. For example, a verb may agree with a plural subject in all ϕ-features or, rather oddly, in singular/feminine, even when the subject (the controller) is masculine plural. Developing an idea first briefly sketched—but ultimately not adopted—by Zabbal (2002), we argue that broken plurals are hybrid nouns. Hybrid nouns have been the topic of much recent research (Corbett, 2000, 2015; den Dikken, 2001; Wechsler and Zlatić, 2003; Danon, 2011, 2013; Matushansky, 2013; Landau, 2015; Smith, 2015): either their syntactic or semantic features can be the target of agreement, creating the possibility of an agreement mismatch. Using Harbour's (2011, 2014) theory of number, coupled with some innovations, we provide the featural make-up of Tunisian Arabic broken plurals and contrast it with that of collectives, on the one hand, and sound plurals, on the other. We propose that the feminine agreement seen with broken plurals is associated with a [+ group] feature, one that is exponed as *-a*. In the course of the discussion, we will argue that all gender features are visible at LF (Hammerly, 2018) and that semantic agreement is routinely possible with nouns that are low on the Animacy Hierarchy.

Keywords

1   Introduction

The aim of this paper is to account for the following puzzle: in Tunisian Arabic (TA, henceforth), the ϕ-features of broken plural subjects normally agree with the verb in gender and number, as in (1a) (masculine plural), but as (1b) shows, agreement can also fail to match. Here, *rjel* 'men' interpreted as masculine plural, unexpectedly triggers feminine singular agreement on the verb (on this type of optional agreement in dialects of Arabic other than Tunisian Arabic, see descriptive work by Wright, 1933; Holes, 1990; Belnap 1991; Brustad, 2000; and formal work by Zabbal, 2002).

(1)   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.'

It has been noticed in the literature on certain dialects of Arabic (Wright, 1933; Brustad, 2000; Zabbal, 2002, and several others) that, when the optional situation in (1) occurs, plural/masculine agreement correlates with a distributive interpretation whereas singular/feminine agreement correlates with a collective reading. The same facts obtain in TA. (1a) receives a distributive interpretation whereas (1b) receives a collective interpretation. In the right contexts, (1a) means that the men went out separately, one by one, but in (1b), that they went out together.

   Only broken plurals are subject to this phenomenon, as evidenced by the ungrammaticality of (2a), where the subject is a masculine sound plural failing to trigger feminine singular agreement on the verb. In Arabic, two plurals exist: a sound plural and a broken plural (Ojeda, 1992; Acquaviva, 2008). Sound plurals are formed by suffixation, while broken plurals are formed by a change in the stem. As shown in (2b), masculine sound plurals can only trigger masculine plural agreement on the verb.

(2)   a. **El   muʕalm-een   rajʕ-et   l-el-biru.*   [TA]
          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 to note at the outset that the "weak" agreement (a pre-theoretical term) exhibited in (1b) is not partial agreement, a well-known phenomenon in Standard Arabic. SVO sentences in Standard Arabic trigger all the ϕ-features of the verb (e.g., 3, MASC, PL), as shown in (3a). On the other hand, in VSO sentences, as in (3b), the verb is inflected in the singular even though the subject noun is plural. (3c) shows that singular marking is obligatory in VSO orders.

(3) a. *Sl-ʔawlaad-u   jaaʔ-uu.*   [Standard Arabic]
      the-boys-NOM came-3.MASC.PL
      'The boys came.'

    b. *Jaaʔ-a   al-ʔawlaad-u.*
      came-3.MASC.SG the-boys.NOM
      'The boys came.'

    c. **Jaaʔ-uu   al-ʔawlaad-u.*
      came-3.MASC.PL the-boys.NOM
      'The boys came.' (Mohammad 1990: 95)

TA is an SVO dialect.[1] This means that the "weak" agreement in (1b) has nothing to do with the phenomenon known as "partial agreement", seen in word order alternations of the type found in Standard Arabic. In addition, not only is (1b) singular, but it is also feminine, showing a gender shift when comparing with (1a) whereas (3b) exhibits no change in gender: the verb is singular, but masculine; a fact that strongly indicates the two phenomena are distinct.

    The phenomenon in (1b) recalls, on the other hand, "deflected" agreement (Ferguson, 1989), as described for Standard Arabic: non-humans and inanimates obligatorily trigger third person feminine singular, as in (4).[2]

---

1 Except for subject nouns in intransitive constructions in which case singular/feminine agreement is possible on the verb. Consider the following example:
   (i) *wesl-et   ktobb-etek.*   [Tunisian Arabic]
     arrived-FEM.SG book.PL-POSS.2.SG
     'Your books arrived.'
2 Kramer and Winchester (2018) focus on this problem in Saudi Arabic. This type of agreement is referred to as "gender switch" and the authors use a Distributed Morphology approach. They

(4)   *ja:ʔa-ti*   *l-kila:bu.*   [Standard Arabic]
      came-FEM.SG the-dog.PL
      'The dogs came.' (Fassi Fehri 1988:119)

For humans in Standard Arabic, only (3a) is possible where full agreement is
seen, while (5) is ungrammatical.[3]

(5)   **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-human and inanimate
subjects can also inflect in the third person feminine/singular, but this is an
optional process, as seen in (6), and importantly, it is also possible with
humans, as seen in (1).[4]

(6)   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.'

Much has been written recently about agreement failures and agreement mis-
matches (Percus 2011; Preminger 2014; Landau 2015, etc.), and the contrast
exhibited in (1) is an interesting case study to add to this literature. Our hypoth-
esis is that the broken plural in (1b) is in fact singular and feminine; the
verb thus agreeing strictly with its subject. We argue that broken plurals are
hybrid nouns. Hybrid nouns have mismatching syntactic/semantic ϕ-features

---

argue that the Saudi Arabic gender and number switch is a syntactic effect because feminine
singular agreement with non-human plural nouns triggers a particular semantic/pragmatic
interpretation, namely, that of a non-individuated herd/clump. Their main claim is that if the
gender switch in question were due to syncretism, it would happen too late in the derivation
(post-syntactically) to affect the semantics. Our proposal is compatible with these ideas.

3   In pre-Islamic poetry and the Qur'an, nonhuman plural controllers did not require feminine
    singular agreement, but feminine singular alternated with full plural agreement, as in TA and
    other modern dialects (D'Anna, 2017).

4   This is a feature of the dialects and is widespread geographically, see Boris (1945–1948), among
    others.

and are thus able to trigger either syntactic or semantic agreement: Landau (2015) and others; Corbett (2000, 2015); den Dikken (2001); Wechsler and Zlatić (2003); Danon (2011, 2013); Johnson and Joseph (2014); Smith (2015). Examples exist in Icelandic, Serbo-Croatian, British and Canadian English (groups nouns, e.g., *committee*), German (*Mädchen*), Hebrew (*be'alim*), polite plurals (Comrie, 1975), pancake sentences (Enger, 2004, 2013), Russian *vrač* 'doctor' (Matushansky, 2013), etc.[5] As proposed by Corbett (2000), controllers of agreement—typically nouns—carry two sets of ϕ-features, syntactic and semantic, which often match, but need not to. Most of the time, it is impossible to tell whether agreement is syntactic or semantic, since the two types of ϕ-features usually coincide. The only way to determine whether the controllee will agree with syntactic or semantic phi-features is by looking at mismatch situations. Focusing on Arabic broken plurals, we propose in a nutshell that, while the agreement is syntactic in (1b), it is semantic in (1a).

In an appendix, Zabbal (2002) proposes a similar idea as an alternative proposal to his main thesis according to which sound plurals denote sums and broken plurals denote sums or groups with sound plurals associated with Num and broken plurals on their group interpretation associated with a lower projection akin to *n*. The alternative proposal is that broken plurals are always associated with a projection closer to N while receiving a uniform semantics (as group plurals). This means that full agreement with the verb is viewed as semantic agreement. Such a proposal immediately raises problems. First, as pointed out by Zabbal himself, it does not explain why the verb is feminine when agreeing in the singular. Second, assuming the broken plural is associated with a feminine morphosyntactic feature, this feature would block the agreement relationship between the noun's semantic features and the verb. This problem can be solved by adopting recent analyses of gender (e.g., Kramer 2009, 2015) where some gender features are interpretable while others are uninterpretable. This would make the feminine form associated with broken plurals uninterpretable while the gender features of the controller would be semantic (natural gender).

However, there is a major problem with this. The fact that (6a) is possible shows that semantic agreement is not necessarily tied to natural gender. The word *biben* 'door', unlike *rjel* 'men', is not associated with natural gender, and thus does not presumably carry an interpretable feature. Yet, semantic agreement is possible between the noun *biben* 'door' and the verb.

---

5 As we shall see, this does not mean all hybrid nouns are similar: they are subject to selectional/lexical parameters, see Landau (2015).

In this paper, we argue that the feminine feature [+FEM], exponed by -*a*, can be associated with a number feature [+group]. In order to explain why it is possible for the gender features of nouns such as *biben* 'door' to be active for semantic agreement, we will propose that all gender features are interpretable as in Hammerly (2018) or that, at least, they are visible at LF.

In sum, we propose that broken plurals, as hybrid nouns, are semantically plural but syntactically singular. When agreeing in the singular, broken plurals are interpreted as groups and the role of the feminine feature [+FEM], exponed by -*a*, is to signal [+group]. We will argue that the number feature [+group], spelled out as feminine agreement, is associated with Num (and ultimately D) and that it does not surface on the broken plural itself (only on agreeing associated lexical items, e.g., verbs, adjectives, that agree with the broken plural), because of the impossibility for broken plurals to have suffixes.

Our article is structured as follows. In Section 2, we introduce Harbour's (2011, 2014) theory of number and the relevant facts about plurals in TA. In this section, unlike Zabbal (2002), we do not propose that sound plurals and broken plurals are generated in two different functional heads (one inflectional, the other derivational). Instead, we propose that both types of plurals are in Num. In Section 3, we summarize Zabbal's (2002) proposal(s) that tackle the problem exhibited by (1a) versus (1b). In Section 4, we give an analysis of broken plurals as hybrid nouns, comparing them with collectives and sound plurals, and we explain how semantic agreement is made possible. We conclude in Section 5.

2 Number in Tunisian Arabic

Our analysis is framed within Harbour's (2011, 2014) feature theory of number. Although the present article focuses on the [+group] feature and the puzzle introduced in Section 1, it is necessary to introduce the details of such a theory, since our analysis of singulars, plurals, and singulatives, relies on it (see Dali and Mathieu 2020). With the analysis of Arabic number in place, we will be able to compare our view of TA plurals with that of Zabbal (2002). We will see that group plurals versus sum plurals do not occupy different functional heads, as proposed by Zabbal (2002), but are instead uniquely associated with Num.

First, we assume, following Harbour (2011, 2014)—and many others—that classificatory features occupy their own projection, namely *n* (= Class) and that *n* takes a root as a complement, as in (7). Furthermore, *n* labels the root as a noun and makes it visible to the computational system.

(7)   nP

    n   √

Second, we assume that *n* defines a nominal predicate P and structures the root as a join semilattice (Harbour, 2011, 2014; Zabbal, 2002; Martí, 2018) giving us the representation in (8) for the semi-lattice.[6]

(8)

$$a + b + c$$

$$a + b \qquad a + c \qquad b + c$$

$$a \qquad\qquad b \qquad\qquad c$$
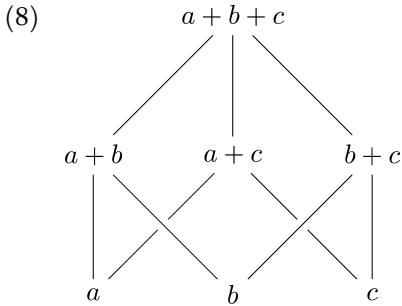
The extended projection of *n* looks like (9) (Borer, 2005; Grimshaw, 2005). NumP takes *nP* as complement and DP takes NumP as complement (Borer 2005, Num = Div).

(9)   DP

   D   NumP

        Num   nP

               n   √
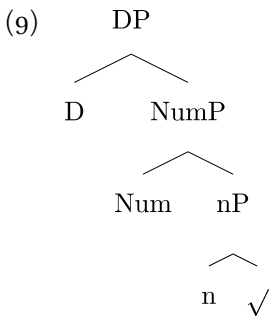
Third, following Harbour (2011, 2014), we assume these nominal functional projections come with features. The features on *n* and Num are semantic in nature (they are interpretable). Those on D are syntactic (uninterpretable):

6   As pointed out by Harbour (2011), *n* underdetermines whether the lattice has an atomic stratum or whether its sub-parts have ever smaller sub-parts, i.e., whether it is count or mass. Like Borer (2005), we assume it is Num (or Div for Borer) that actually introduces the distinction. When Num is projected, the noun is count, when Num is not projected, the noun is mass.

they encode number agreement. Number is determined by cardinality (singular, dual, plural), and *n*, by semantic properties of the noun. The semantic features for number features are as follows (see also Noyer 1992):

(10) a. [+atomic]   (11) a. [+minimal]   (12) a. [+additive]
     b. [−atomic]   b. [−minimal]   b. [−additive]

As summarized by Martí (2018:7): "[±Atomic] is sensitive to atoms/non-pluralities ([+atomic]) vs. non-atoms/pluralities ([−atomic]). [±Minimal] is sensitive to elements with parts ([−minimal]) vs. elements without parts ([+minimal]). [±Additive] is concerned with whether the output set contains, for any two of its members, their join ([+additive]) (a property also known as cumulativity; cf. Krifka 1989) or not ([−additive])".[7]

The features above allow us to account for singulars, plurals, and duals in TA. However, in order to account for collectives in TA, we need the following set of features. As a (non-semantic) classificatory feature, [+collective] appears on *n* (this feature is an innovation and not found in Harbour's 2011, 2016 theory).

(13) [+collective]
     [−collective]

Now, we turn to Tunisian Arabic and lay out the structures for singulars, plurals, and collective nouns. A singular noun, *kalb* 'dog' (14a), masculine non-human or *fannen*/*fannena* 'artist' (14b), masculine/feminine human, has the structure in (15). Num, since it is here singular, is associated with the features [+atomic;

---

7   It must be noted that not all features are available in all languages. For example, not all languages have a paucal, a dual, a greater plural, etc. Variation is expressed by (i).
    (i)   Activation parameter: [+/−additive]/[+/−minimal]/[+/−atomic] is (not) a feature of Number. (Harbour 2014. p. 203)
    For example, "in order for a language to have an approximate number, this parameter must be active (Languages, like English, in which it is inactive, have no approximate numbers.)" (Harbour 2014, p. 198).
        It must also be noted that social convention intervenes in the interpretation of [±additive] and that this constitutes further variation between languages. For example, TA has ten as an upper-bound cutoff while Bayso (Cushitic) has six (Corbett, 2000).
    (ii)   Sociosemantic convention: The semantic range of the cut defined by [±additive] is subject to social convention. (Harbour 2014, p. 198)
    Finally, it is possible (or impossible, depending on the language) for features to be recursive, as expressed by (iii).
   (iii)   Feature recursion parameter: both values of [±F] may (not) cooccur on Number⁰. (Features so parametrized are starred, [±]*.) (Harbour 2014, p. 203)

–additive]. In this article, we do not make use of the feature [+minimal], since the use of the dual is sparse in TA (see details in Blanc, 1970). In Standard Arabic, Num would have an extra feature, namely [+minimal]. Note that *fannen-fannen-a* is a contrasting pair, and thus the morpheme *-a* obligatorily marks the feminine gender.

(14) a. *kalb*   [Tunisian Arabic]
        dog.SG
        'a dog'

    b. *fannen*   *fannen-a*
        artist.MASC.SG   artist-FEM.SG
        'a (male) artist, a (female) artist'

(15)
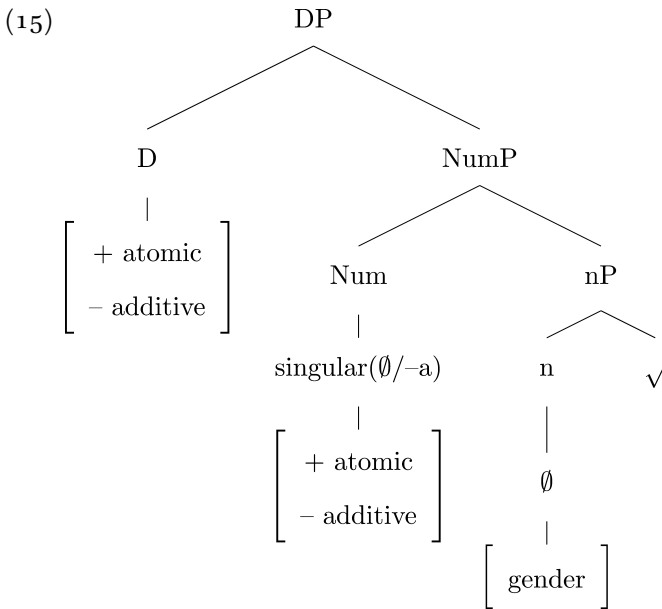
In (15), *n* is unmarked for number (Harbour, 2011, 2014). This is indicated by ∅. In contrast, collective nouns are marked [+collective] (see below). Num, since it is here singular, is associated with the features [+atomic; –additive]. The features on Num are repeated on D to indicate agreement (e.g., with the verb). Finally, we assume that gender features are associated with *n* (Kramer, 2009, 2015; Kihm, 2003; Lowenstamm, 2008): feminine or masculine depending on the noun. Singular masculine is spelled out as ∅ while singular feminine is spelled out as *-a*.

The plural in TA comes in two shapes. The sound plural in (16) is formed by suffixation of a plural marker on an unmodified stem. The suffix *-een* marks the plural of masculine nouns (16a), while *-et/-at* marks the plural of feminine nouns (16b).

(16) a. *fannen*   → *fannen-een*   [Tunisian Arabic]
       artist.MASC.SG   artist-MASC.PL
       'artist, artists'

      b. *fannen-a*   → *fannena-at*
       artist-FEM.SG   artist-FEM.PL
       '(female) artist, (female) artists'

The broken plural is formed by stem change (17). No specific suffixal morpheme in the broken plural is associated with the plural meaning.

(17) *kalb*   → *kleb*   [Tunisian Arabic]
      dog.SG   dog.PL
      'dog, dogs'

Following Dali and Mathieu (2020), we propose that both sound plurals and broken plurals are associated with Num, in which case we have (18). We will give the features sound and broken plurals are associated with in Section 4 (the features they are not associated with are not identical).[8]
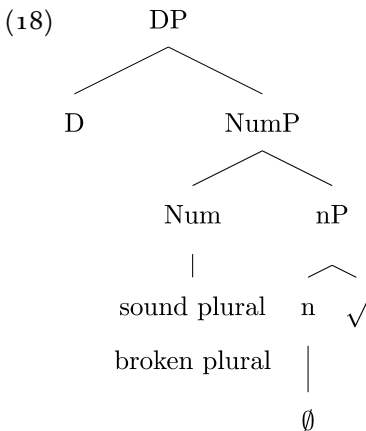
(18)

This differs from the account proposed by Lahrouchi and Lampitelli (2015), based on data from Moroccan Arabic. They argue that the broken plural is realized on the *n* head, whereas the sound plural realizes a feature on the number head. The motivation behind their *n*-based analysis of the broken plural is that, in their view, internal (broken) plurals behave irregularly (i.e., they undergo a variety of morphophonological changes such as infixation, vowel insertion and gemination). In their analysis, the external (sound) plurals, on the other hand, invariably use the suffix -*a*. They also claim that, from a semantic point of view, internal plurals are often semantically associated with collective readings. For Lahrouchi and Lampitelli (2015), the distinct syntactic locations of these plurals allow one to capture their morphophonological and semantic differences. Their account is in line with previous split plurality accounts (Lecarme, 2002 for Somali; Kramer, Kramer (2016) for Amharic). As we will argue in Section 3 (see also Dali and Mathieu 2020), everything points to the view that broken plurals are not irregular forms. In addition, both sound and broken plurals can refer to sums and thus both must be operations on semi-lattices, which means they are identical semantically (and when the broken plural is interpreted as a group, it still operates on a semi-lattice).

In addition to the singular/plural contrast described above, most Semitic languages also have a singulative system, where the collective is the default number from which the singulative is derived. For Semitic languages, the collective-singulative distinction is most productive in Arabic, Maltese, and Ethio-Semitic languages—with Hebrew having retained only residues of the system (Doron and Müller, 2013). The following examples illustrate the singulative system of Arabic (19). The left column features collective nouns while the right column exemplifies singulative forms.

(19) a. *ðahb*   ~ *ðahb-a*   [Tunisian Arabic]
        gold.MASC.COLL   gold-FEM.SING
        'gold, a bit/nugget of gold'

     b. *jormen*   ~ *jormen-a*
        duck.MASC.COLL   duck-FEM.SING
        'ducks, a duck'

     c. *dud*   ~ *dud-a*
        worm.MASC.COLL   worm-FEM.SING
        'worms, a worm'

The input nouns to the singulative (the collectives on the left in (19)) are always masculine in Tunisian Arabic and Maltese, and although they refer semantically to sums, they trigger singular agreement on verbs and other dependent categories. Consider (20). The plural subject *nemel* 'ants' agrees with the verb in third person masculine singular.

(20) *Nemel*   *dxal*   *l-el*   *kujina.*   [Tunisian Arabic]
     ant.MASC.COLL entered.3.MASC.SG to-the kitchen
     'Ants entered the kitchen.'

The singulative is derived from the collective through suffixation of the marker *-a* (on the singulative in Arabic, see Ojeda 1992; Zabbal 2002; Fassi Fehri 2003, 2012; Borer and Ouwayda 2010; Mathieu 2012, 2009, 2014). Unlike collectives, singulative nouns are feminine. This shows both morphologically and syntactically.
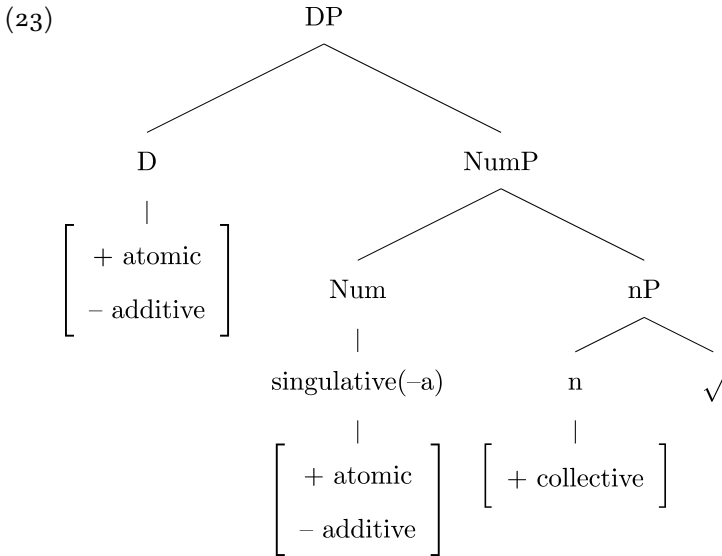
First, the ending, *-a*, is a typical feminine ending in Arabic and Maltese, and can be found in morphological feminine nouns such as semantically feminine nouns (21).

(21) *ħajjem*   ~ *ħajjem-a*
     hairdresser.MASC.SG   hairdresser-FEM.SG
     '(male) hairdresser, (female) hairdresser'

Second, singulative nouns in Arabic (and Maltese) consistently trigger feminine agreement on the lexical items they control. Consider (22) and compare this example with (20). While in (20) the collective noun *nemel* 'ants', a morphosyntactically masculine singular form, triggers masculine singular agreement on the verb, its singulative counterpart *nemela* 'an ant', in (22), triggers feminine singular agreement on the verb.

(22) *Nemel-a*   *daxl-et*   *l-el*   *kujina.*   [Tunisian Arabic]
     ant-FEM.SING entered.3-FEM.SG to-the kitchen
     'An ant entered the kitchen.'

We see that the singulative chooses a unit from a collection or a portion from a mass term. We propose (23) as the representation of a singulative operating on a collective noun.

(23)

```
                              DP
                   ╱                    ╲
                  ╱                      ╲
                 D                       NumP
                 │                 ╱            ╲
          ┌─ + atomic  ─┐        ╱              ╲
          │             │      Num               nP
          └─ − additive ─┘      │           ╱        ╲
                         singulative(−a)    n          √
                               │            │
                        ┌─ + atomic  ─┐  ┌ + collective ┐
                        │             │  └              ┘
                        └─ − additive ─┘
```

We introduce the singulative for two reasons: 1) it is a way to introduce collective nouns in Arabic and introduce the [+collective] feature they are associated with; 2) it is a way to introduce the *-a* morpheme, which is used not only for feminine nouns in TA, but a variety of uses, including the singulative, but also group nouns, as we shall see in Section 4. Note that the input to the singulative operation is a [+collective] noun rather than a count noun (marked ∅ above).[9]

To summarize Section 2: we introduced Harbour's (2011, 2014) theory of number and gave structures for singulars, plurals, singulatives, and collectives in TA. Before we turn to our analysis of the contrast in (1), we introduce Zabbal's (2002) theory of Arabic plurals.

---

9   *-a* in Arabic is thus ambiguous. It has different denotations depending on the base noun being used. When it is used with a collective noun, it derives a singulative. When it is used with a count noun, it derives a group. Historically, these uses are related, see Section 4 for discussion. In other words, [+FEM], exponed by *-a*, creates a kind of polarity effect (see Grimm, 2012 on *-ri* in Dagaare that can be used both as a plural marker or as a singulative marker depending on the base).

3   Zabbal (2002)

Zabbal (2002) makes a distinction between s-plurals (sum plurals) and g-plurals (group plurals). Sound plurals are s-plurals: they refer to sums or sets. Broken plurals can either be s-plurals or g-plurals. This is compatible with a view that broken plurals are ambiguous (rather than hybrid nouns): they constitute two discrete forms instead of one, the latter solution being clearly preferable in that it satisfies Occam's razor. On Zabbal's view, broken plurals are s-plurals when the verb carries plural agreement, but they are g-plurals when the verb carries singular/feminine agreement. Sound plurals are always s-plurals, since they always denote sums and agree with the verb in the plural. Zabbal argues that the g-plural is associated with N (making it lexical and derivational) while the s-plural is under Num (inflectional). This is represented in (24).

(24)
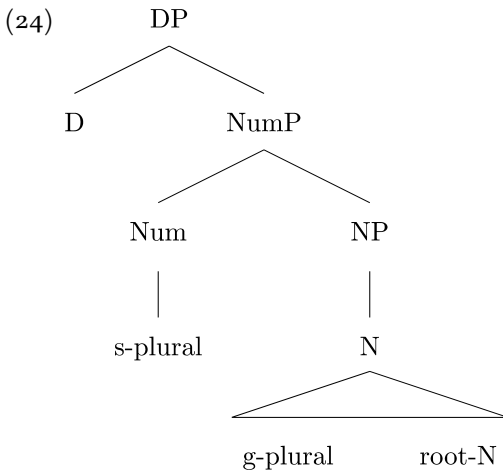
There is a major problem with (24). On this view, while s-plurals are inflectional (syntactic), g-plurals are derivational (lexical). An important fact to consider, however, is that broken plural forms are predictable and productive. McCarthy and Prince (1990) note that the broken plural process is so productive that it easily assimilates neologisms loan words. Consider, for instance, the French borrowing *bank* 'bank' (from the French *banque*), which takes the broken plural *bunuk* 'banks'. Therefore, we do not consider them as lexical.

In addition to productivity, many arguments support an inflectional analysis of the broken plural. First, there is a clear and systematic prosodic connection between broken plurals and their respective singulars, which makes them much less arbitrary than English irregular plurals, e.g., *book/book-s* vs.

woman/women (see Acquaviva 2008, McCarthy and Prince 1990, Ratcliffe 1998, Kihm 2003). Arabic has canonical stems from which broken plurals can be created, and broken plurals form 70% of plural nouns in Arabic and are often, if not always, the default plural form (contrary to general beliefs that plural suffixes are the default).
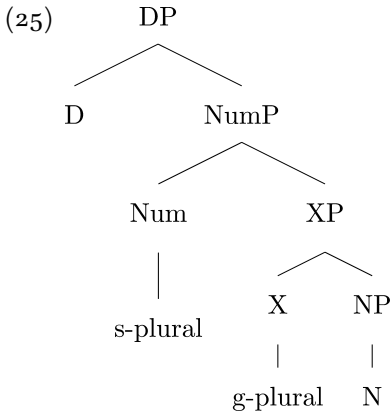
Second, some accounts of the broken plural (Ratcliffe, 1998; Kihm, 2003; Lahrouchi and Lampitelli, 2015) treat it like an operation on *n*, where the broken plural vocalic pattern is applied to the root to give a noun, as an alternative to the singular noun. However, it appears that broken plurals are operations on a stem, and not on a root. For example, the properties of the singular stem are carried over to its broken plural counterpart. These properties include vowel quantity, number of syllables, and consonant spreading (McCarthy and Prince, 1990; Hammond, 1988). The root itself contains no such information (the Semitic root consists only of consonants). Therefore, several studies of the Arabic broken plural established the form of the singular as the principal factor determining the form of the broken plural (Murtonen, 1964; Levy, 1971).

Third, broken plural patterns also apply to adjectives (e.g., *mrið/morða* 'sick'). Considering this, a stem-forming process analysis is very unlikely. Adjectives typically do not get number information as part of their derivational formation (which is attributed to the head 'a(djective)' in the DM framework), but rather acquire this inflectional information through agreement. Claiming that the broken plural is on *n* does not account for this fact.

Based on these arguments, there is no reason to treat g-plurals as derivational or lexical phenomena.[10]

Aware of these problems, Zabbal (2002) proposes a brief alternative analysis towards the end of his thesis: the g-plural is in fact inflectional and thus not under N. 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 (25).

---

10   According to Acquaviva (2008), broken plurals are lexical in that they are stem forms (produced via Level 1 morphology), but inflectional in that they express number information (via Level 2 morphology). On his view, the broken plurals are in *n* and a morphologically null Number head appears above to express the broken plural's inflectional properties. This higher dividing operator is necessary for syntax, but not for morphology (Acquaviva 2008:271). On our view, broken plurals are not in *n*, only in Num.

(25)

```
                DP
              /    \
            D       NumP
                   /    \
                 Num      XP
                  |      /   \
              s-plural  X     NP
                        |      |
                    g-plural   N
```

As Zabbal (2002) himself realizes, there is still a major problem with this pro-posal in that, although the sum broken plural and the group broken plural have the same morphological form, they still appear on two different syntactic nodes. If we adopt the view according to which complementary distribution is the hallmark of identity and that this is reflected in the nominal structure, we should expect s-plurals and g-plurals to be generated under the same node. In addition, the difference between broken and sound plurals appears mainly morpho-phonological. Thus, why should they appear on different nodes?

Another problem is that, while sound plurals are always s-plurals, broken plurals are sometimes s-plurals (when they agree with the verb in the plural) and sometimes g-plurals (when they agree with the verb in the singular). In other words, depending on the interpretation, broken plurals have not only a different syntax (as seen above), but a different semantics, depending on whether they agree in the singular or the plural.

Not entirely happy, it seems, with these proposals, Zabbal (2002) entertains the idea, in a short appendix (Section 4.9), that broken plurals have a unified semantics. Zabbal notes a similarity between broken plurals in Arabic and com-mittee nouns in English, and suggests that while all broken plurals are g-plurals with one semantics (group plurals), the distributive reading and the plural agreement on the verb are manifestation of purely semantic number. However, Zabbal (2002) is not sure this analysis will work because it has, according to him, difficulties in accounting for the fact that there is a difference in gender between the g-plural and the s-plural. The g-plural always has feminine gender and the s-plural has the gender of its underlying noun. The problem seems to be the following: assuming that gender features are morphosyntactic features, the feminine form associated with the broken plural intervenes with the agree-ment of the features of the noun with the verb.

In the next section, we build on Zabbal's (2002) insight, providing an account that solves this problem. Our account will also solve the problem associated with (6). Nouns that carry natural gender are not the only nouns that can trigger feminine singular agreement on the verb, nouns that carry arbitrary gender can do the same. This is completely unexpected under a simple semantic agreement account.

4 Broken plurals are hybrid nouns

The unusual pattern of agreement of TA introduced in (1b) is reminiscent of the behaviour of hybrid nouns. As pointed out in the introduction, hybrid nouns have mismatching syntactic/semantic ϕ-features and are thus able to trigger either syntactic or semantic agreement (Landau, 2015 and others, Corbett 2000; 2015, Den Dikken, 2001; Wechsler, 2003; Danon, 2011; 2013; Matushansky 2013; Smith, 2015).

Let us take, as a way of illustration, 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 relevant examples in (26). It is possible for the verb to be inflected in the singular (26a) or in the plural (26b).

(26) a. The committee is meeting next week.
      b. The committee are meeting next week.

The idea is that group nouns have the feature matrix in (27) where either the syntactic features or the semantic features can be accessed. In (26a), the syntactic features are activated whereas in (26b), the semantic features are at play. Like others (den Dikken, 2001; Gardelle, 2019), we assume that the singular is the default agreement pattern and that the plural agreement pattern is an added/superimposed effect (possibly of the universal Animacy hierarchy, Gardelle, 2019).

(27) $\begin{bmatrix} \text{syn: } 3 & \text{SG} \\ \text{sem: } 3 & \text{PL} \end{bmatrix}$

This contrasts with plural subjects, since they have the feature grid in (28).

(28) $\begin{bmatrix} \text{syn: } 3 & \text{PL} \\ \text{sem: } 3 & \text{PL} \end{bmatrix}$

It has been noticed that the contrast in agreement observed with group nouns correlates with a distributive / collective distinction. For example, in (29a) the predicate *is old* is predicated of the committee as a whole. It is the committee that is old and not necessarily its members. This is the collective reading. In (29b), on the other hand, the predicate *are old* is predicated of the members of the committee. It is the members of the committee that are old and not necessarily the committee itself. This is the distributive reading (example from Barker, 1992, p. 89).

(29) a. The committee is old.   (collective)
     b. The committee are old.   (distributive)

Predicates such as *be old* are usually distributive. This can be seen in (30) where a plural subject necessarily triggers plural agreement and is thus interpreted as distributive. The collective reading is not available (the men cannot be old together, 'be old' is an individual property), and **the men is old* is not grammatical.

(30) The men are old.   (distributive)

In (29a), the agreement is with the syntactic features of the group noun, but in (29b), the verb agrees with the semantic features of the group noun.

In the case of collective predicates, we have similar results. Singular agreement correlates with a collective reading (31a) while plural agreement correlates with the distributive interpretation (31b).

(31) a. The committee is gathering.   (collective)
     b. The committee are gathering.   (distributive)

With plural subjects and collective predicates, only the collective reading is available, and singular inflection is impossible **The men is gathering*.

(32) The men are gathering soon.

Of course, ambiguous predicates such as *leave* will allow both options. This is true for committee nouns (33) and plural subjects (34).

(33) a. The committee is leaving.   (collective)
     b. The committee are leaving.   (distributive)

(34) The men are leaving.   (collective, distributive)

In sum, English plural subjects and committee nouns behave differently: only
plural agreement is possible with plural subjects, as summarized by the exam-
ples shown in (35), (36), and (37). Note that, since a plural subject is used, (35a)
is strictly distributive, (36a) strictly collective, and (37a) is, of course, ambigu-
ous.

(35) a. Men are coughing.
      b. *Men is coughing.

(36) a. Men are gathering.
      b. *Men is gathering.

(37) a. Men are leaving.
      b. *Men is leaving.

To recapitulate, group nouns are special in that they can agree with the verb
in the singular (and thus yield a collective reading) or in the plural (and yield a
distributive reading) no matter which kind of predicate is used while, of course,
plural subjects are restricted, since they refer both syntactically and semanti-
cally to pluralities. What we see with group nouns is that verb agreement marks
the collective/distributive distinction (de Vries, 2013, 2015).

   Turning now to TA, we see that broken plurals behave like British/Cana-
dian English group nouns rather than plural subjects. Broken plural subjects
in Tunisian Arabic can agree in the singular or in the plural quite freely: with
distributive (38), collective (39), as well as ambiguous predicates (40).

(38) a. *rjel   i-koħ-u.*   (distributive)
          man.PL 3.MASC-coughing-PL
          '(Some) men are coughing.'

      b. *rjel   t-koħ.*   (collective)
          man.PL 3.FEM-coughing.SG
          '(Some) men are coughing.'

(39) a. *El   rjel   tlamm-u.*   (distributive)
          the man.PL gathered-3.MASC.PL
          'The men gathered.'

b. *El   rjel   tlamm-et.*   (collective)
   the man.PL gathered-3.FEM.SG
   'The men gathered.'

(40) 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.'

Sound plurals, on the other hand, behave like normal plural subjects. Consider the following examples. (41a), (42a), and (43a) are all grammatical, but (41b), (42b), and (43b) are not. As pointed out for British English plurals above, (41a) is distributive, (42a) is collective, and (43a) is ambiguous.

(41) a. *muhands-een   i-koħ-u.*   (distributive)
        engineer-MASC.PL 3.MASC-coughing-PL
        '(Some) engineers are coughing.'

   b. **muhands-een   t-koħ.*
       engineer-MASC.PL 3.FEM-coughing.SG
       '(Some) men are coughing.'

(42) a. *El muhands-een   tlamm-u.*   (collective)
        the engineer-MASC.PL gathered-3.MASC.PL
        'The engineers gathered.'

   b. **El   muhands-een   tlamm-et.*
       the engineer-MASC.PL gathered-3.FEM.SG
       'The engineers gathered.'

(43) a. *El muhands-een   xerj-u.*   (distributive/collective)
        the engineer-MASC.PL went.out-3.MASC.PL
        'The men went out.'

   b. **El   muhands-een   xerj-et.*
       the engineer-MASC.PL went.out-3.FEM.SG
       'The engineers went out.'

These are the sorts of facts that lead us to propose that broken plurals in TA are hybrid nouns. The idea is that broken plurals are syntactically singular, and so they strictly agree with the verb or the adjective in the singular. As in the case of group nouns described above, this is the default agreement pattern. When the agreement on the verb or the adjective is plural, it is the semantics features on Num that are accessed. This is a superimposed interpretive effect that overrules the default agreement setting. (44) is the feature grid for broken plurals.

(44) $\begin{bmatrix} \text{syn: 3} & \text{SG} \\ \text{sem: 3} & \text{PL} \end{bmatrix}$

This is interpreted as a group. Let us assume that a group operator is generated in Num and that it turns a plural NP into an atom. We propose further that -*a* expones a feature [+FEM] and that this morphosyntactic feature can express [+group], as in (45).

(45) [+FEM] = [+group]

This is not unlike what happens with the singulative feminine, except that the interpretation is the reverse and that the exponent is not spelled out on the broken plural itself. This is because broken plurals are formed by a change in the stem, and gender can generally only be expressed through suffixation in Arabic (Caubet, Simeone-Senelle, Vanhove 1989, Wright 1933:183). However, the [+group] feature expressed by -*a* surfaces on D and as such agrees with verbs and adjectives. Below, we will provide arguments in favour of the view that -*a* can express groups.

(46), on the other hand, gives the feature grid for sound plurals. Such plurals are syntactically plural and as such agree with the verb only in the plural.

(46) $\begin{bmatrix} \text{syn: 3} & \text{PL} \\ \text{sem: 3} & \text{PL} \end{bmatrix}$

Turning now to gender features of such plurals, we adopt the recent analyses of gender (e.g., Kramer 2015) that have proposed some gender features are interpretable while others are uninterpretable and that both types are generated in *n*. This distinction is equivalent to the more traditional distinction between natural versus arbitrary gender (Corbett, 1991). (47) is Kramer's (2015) feature structure.

(47) Kramer (2015) feature structure
    a. Arbitrary masculine: n [ ]¹¹
    b. Arbitrary feminine: n [u +fem]
    c. Natural masculine: n [i –fem]
    d. Natural feminine: n [i +fem]

The syn(tactic) features in the feature grids above are equivalent to uninterpretable features and the sem(antic) features are equivalent to interpretable features.

    The question that now arises is the following: How do we account for semantic agreement in a theory with AGREE? We follow Smith (2015, 2017)—see also Arregi and Nevins (2012)—in assuming a modified version of AGREE where the operation proceeds in two steps, as follows (48) (Smith 2017, p. 18).

(48) Agreement by Probe with Goal proceeds in two steps:
    a. AGREE-LINK: a probe has unvalued ϕ-features that trigger Agree with a goal (possibly more than one). The result is a link between probe and goal.
    b. AGREE-COPY: After the syntactic derivation, the values of the ϕ- features of the goal are copied onto probe linked to it by AGREE-LINK.
  i.   if AGREE-COPY happens at the point of transfer, this requires that goal c-command the probe.

Once the second step (COPY) has occurred, the valued features are no longer active. Semantically motivated agreement involves AGREE-LINK, but requires AGREE-COPY to copy feature values from a valued feature. This is made possible because AGREE-COPY happens at the point of Transfer, when both unvalued and valued features are present, leaving the valued features accessible for (semantic) agreement.

    So far, so good. There nevertheless remains an additional problem to be solved. To see the problem, compare (1a), repeated here as (49), and (6a), repeated here as (50).

(49) *El rjel   xerj-u.*   [Tunisian Arabic]
    the man.MASC.PL went.out-3.MASC.PL
    'The men went out.'

---

11   Masculine noun-sex differentiable nouns lack gender features in this system.

(50) *El biben tsakkr-u.*
      the door.PL closed-MASC.PL
      'The doors closed.'

(49) is unproblematic: assuming semantics features are accessed in this case, the gender features associated with 'men' are interpretable. On the other hand, (50) is a priori problematic: this is because the gender associated with 'doors' are normally uninterpretable, and it is not clear how these would be accessed, since they are not semantic features. The word *biben* 'door', unlike *rjel* 'men', is not associated with natural gender, and should technically not carry an interpretable feature. Yet, semantic agreement is possible between the noun *biben* 'door' and the verb.

    To solve this problem, we adopt the view that *all* gender features are interpretable (Hammerly, 2018) or that, at least, they are all visible at LF. Hammerly (2018) builds his theory on examples such as the following.

(51) a. bassin (MASC)   [French]
        'basin (geographical)'
    b. bassine (FEM)
        'washing basin'

(52) a. chapelet (MASC)
        'prayer beads'
    b. chapelle (FEM)
        'prayer room'

(53) a. rouet (MASC)
        'spinning wheel'
    b. roue (FEM)
        'wheel'

(54) a. cache (MASC)
        'mask'
    b. cache (FEM)
        'hiding place'

The idea is that even inanimate/non-natural items can involve gender differences with semantic import. Assuming the pairs above all involve nouns that stem from the same respective roots, we see that a change of gender signals a difference in meaning. There is often an associated change in morphophono-

logical form (e.g., (51)–(53)), but this is not always the case (e.g., (54)). As pointed out by Hammerly (2018), alternations in gender within nouns that denote inanimate objects are far less discussed in the literature, although there are many examples of this process. For the sake of simplicity, we will continue using the terms interpretable/uninterpretable, but we will view both types as visible at LF.

In sum, while hybrid agreement for English collective nouns "is common for pluralities of humans (e.g., *committee*), rarer for animals (e.g., *herd*), and impossible for inanimates (e.g., *the forest are …*)" (Gardelle, 2019, p. 33), hybrid agreement in Arabic is different: it permits agreement with inanimates, i.e., nouns that are basically low on the Animacy Hierarchy (Corbett, 2000; Croft, 2003).

(55) The Animacy Hierarchy
   human > animates > inanimates (Croft, 2003, p. 130)

This situation is unique to Arabic broken plurals. In other languages and/or for other nouns, the semantic override agreement superimposed on the basic agreement pattern is usually, if not always, correlated with items in the higher domain of the Animacy Hierarchy. For example, the German word *Mädchen* *'girl'* is grammatically neuter, and as such, it triggers neutral agreement in all its targets, while the semantic override may trigger the feminine in the personal pronoun (*sie*). In Russian, *vrač* 'doctor' is morphosyntactically masculine, but can be used with a feminine pronoun if referring to a female doctor. Again, we see here a case where the semantic override has to do with animates and not the other way around. However, it must be noted that, since hybrid agreement is dependent on the Animacy Hierarchy, it does not rule out semantic override with inanimates, since the latter are part of, and not excluded by, the hierarchy.

In other words, hybrid nouns are not created equals cross-linguistically, their internal features might not be the same (see footnote 13 below) and not necessarily accessible in the same way. Semantic override works differently for different languages and different constructions/hybrid nouns. Hybrid agreement for both gender and number is governed by two universal hierarchies: a semantic one, the Animacy Hierarchy, and a formal one, the Agreement hierarchy, but how these are manifested depends on the language and the noun in question. A formal account of this is needed, for sure, but this is beyond the scope of this paper.

Let us now give full structures for broken plurals and their associated features (we ignore Person features). First, we assume that sound plurals always denote sums in Arabic and that broken plurals basically denote groups (Zabbal,

2002) (when plural agreement is triggered the interpretation is via semantic agreement). (56) is the structure for a broken plural referring to "men" interpreted as a group. The features on Num are semantic while the features on D are syntactic ($\varnothing$ on *n* indicates the noun class; in this case, a regular count noun, not a collective noun). The [+group] feature is spelled out as -*a* (although not pronounce because gender in Arabic can only be expressed through suffixes in Arabic and not through stem change, as already discussed) and this is picked up in Num. The [+group] feature is similar to [+atomic] in that it can be both a semantic feature (on Num) and a syntactic feature (on D).

(56)
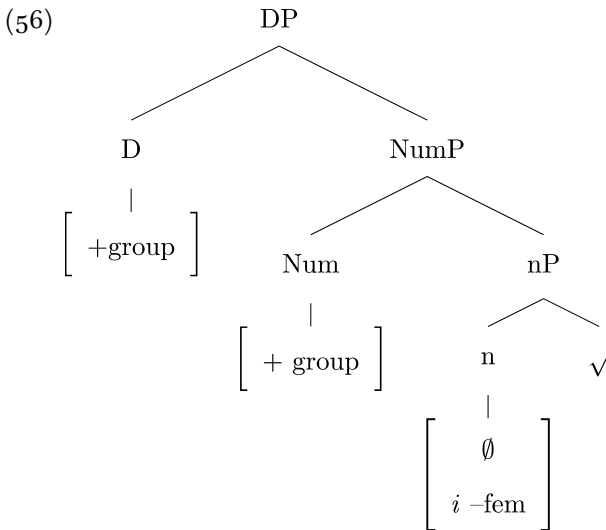
Our account differs from that of Fassi Fehri's (2018) who refers to the feminine broken plurals with a group interpretation as "pluratives". Fassi Fehri claims that the plurative is a special kind of plural that differs in many aspects from the regular broken plural, and hence should receive a distinct syntactic treatment. Because pluratives are both numbered and gendered Fassi Fehri proposes a hybrid structure to express the dual nature of the (Gen/group) head involved. On his view, the plurative is hosted on a Gen(der) head that is dominated by the Num head. The plurative is not just Gen, since it cannot be interpreted as "female", and it is also not just any Num, because although it is a plurality, it cannot be distributive (nor exclusive or inclusive), and it does not control "regular" plural agreement. Taking these two observations into account, Fassi Fehri thus integrates the plurative into a hybrid structure comprising both projections.

On our view, broken plurals are not ambiguous, but only have the structure in (56). The [+group] feature on Num indicates that we are dealing with a group.

As a group, a broken plural agrees syntactically with the verb or adjective in the singular, but since a group contains members and the sum of its members is a plurality, it is possible for semantic agreement to occur. Assume that *n* defines a nominal predicate P and structures the root as a join semilattice (Harbour, 2011, 2014; Zabbal, 2002; Martí, 2018), the group selects for example a group of men from the sums in the semi-lattice.[12]

In short, we have formalized Wright's (1933) original insights, summarized in this quote.

(57) 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)

Collectives and broken plurals are not associated with the same gender (collectives are masculines, broken plurals are feminine). This difference resides in the fact that they do not belong to the same class. While broken plurals are part of the count class, collectives are part of the collective/singulative class. The latter is similar to mass nouns in its distribution: collective nouns, unlike count, cannot be pluralized directly (that is, prior to the singulative operation) (59a) and cannot combine with numerals (59b). In addition, unlike broken plurals, collectives can never trigger semantic (plural) agreement on dependent categories. The only agreement option for collective nouns is masculine/singular (59c).

(58) a. *qattous* → *qtatess* [Tunisian Arabic]
    cat.MASC.SG   cat.PL
    'cat, cats'

   b. *xamsa qtatess*
    five   cat.PL
    'five cats'

---

12   The group feature can be thought of as an operator and made to correspond to Zabbal's (2002:64) definition: $[GP(P)] = x \in At \mid$ there is a $y \in [PL(P)]$ such that $f(x) = y$.

    c. *qtatess y-etʕark-u/t-etʕar-ek*
       cat.PL MASC-fighting-PL/FEM-fighting-SG
       'Cats are fighting.'

(59) a. *djej*   → **dejeyej*   [Tunisian Arabic]
       chicken.MASC.COLL   chicken.PL
       'chickens, chickens'

     b. **xamsa djej*
        five   chicken.MASC.COLL
        'five chickens'

     c. *djej*   *y-etʕar-ek/*y-etʕar-u*
       chicken.MASC.COLL MASC-fighting-SG/MASC-fighting-SG
       'Chickens are fighting.'

Despite the fact they belong to different classes, on our view, broken plurals are similar to collective nouns in that both are semantically plural, but morphosyntactically singular. There is clear evidence that broken plurals are singular.

    First, Acquaviva (2008) observes that some patterns used for the plural of some nouns appear in the singular of other nouns. For instance, the plural *kilaab* 'dogs' has the same prosodic structure and vowel melody as the singular *kitaab* 'book' (plural *kutub*). Therefore, nothing in the morphology of broken plurals indicates that they are *plurals*. Sound plurals, on the other hand, are associated with true plural morphology, namely a plural suffix that is added to the singular shape. This is why sound plurals are never subject to the agreement pattern alternation: they are true plurals.

    Second, it is possible to pluralize broken plurals in TA, as seen in (60) and (61).

(60) *ħsan*   *ħsonna*   *ħsonn-et*   [Tunisian Arabic]
     horse.SING horse.BPL horse.BPL-FEM.SPL
     'horse, horses, horses'

(61) *kteb*   *ktob*   *ktobb-et*
     book.SING book.BPL book.BPL-FEM.SPL
     'book, books, books'

The fact that broken plurals can be pluralized shows that they are singular to begin with. (62) is a pluralized broken plural in a sentence. The plural of the broken plural is interpreted as a pile of books in this context, this shows that we have a plural of a group (a pile can be seen as a group).

(62) *l-bit   l-kolli   ktobb-et   ktobb-et.*   [Tunisian Arabic]
     the-room the-all book.BPL-FEM.SPL book.BPL-FEM.SPL
     'There were piles of books everywhere in the room.'

In Dali and Mathieu (2020), we propose that this kind of pluralization is made possible in Arabic, because renominalization applies. The NumP is turned into another *n*, and NumP applies again. The structure we propose is inspired by Harbour's (2014, p. 221) structure for plurals of plurals, as represented in (63).

(63)
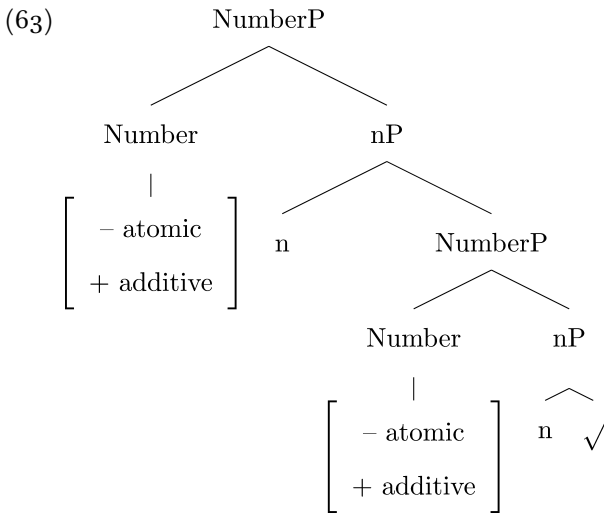
The second *n* introduces a new semi-lattice on which the plural can operate. This proposal explains the residual derivational nature of broken plurals. Derivational accounts might simply argue that broken plurals are in *n* and thus pluralization is achieved via Num. But we reviewed above many arguments in favour of the view that broken plurals are not derivational.

Sound plurals, on the other hand, cannot be pluralized (64). This is because they are plurals syntactically and their featural grid is different from that from broken plurals.

(64) *mrey-et-et*   [Tunisian Arabic]
      mirror-FEM.PL-FEM.PL
      'mirrors'

(65) is the structure for sound plurals. These are interpreted as sums and unlike
broken plurals are not marked [+group].

(65)

```
                              DP
                 ╱────────────────────────╲
               D                          NumP
               │                      ╱─────────╲
        ⎡ − atomic  ⎤         Num(−een/−at)      nP
        ⎢           ⎥              │            ╱╲
        ⎣ + additive ⎦      ⎡ − atomic  ⎤    n    √
                            ⎢           ⎥    │
                            ⎣ + additive ⎦   ∅
```
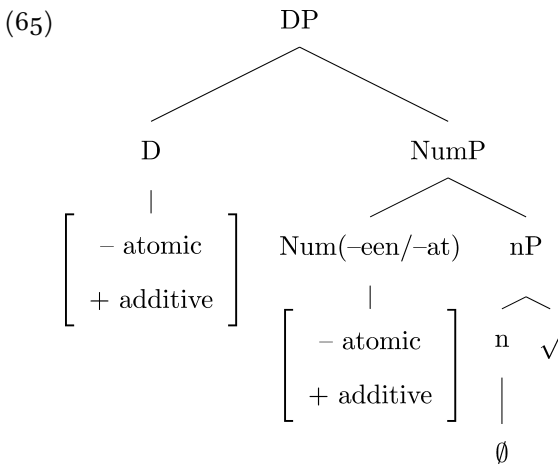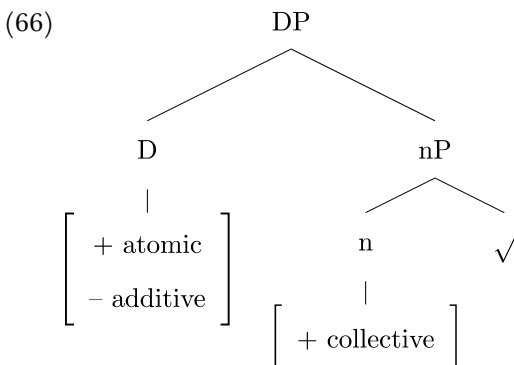
For collectives, we propose the structure in (66), where the [+collective] class
feature appears on *n*. This structure has no Num projection, since it cannot
combine with numerals. Essentially, following what is proposed in Mathieu
(2012, 2009, 2014) and Borer and Ouwayda (2010) as well as Ouwayda (2014), we
assume that the role of the Num functional projection (the equivalent of Div
in Borer's 2005 proposal) is to divide nouns prior to their combination with the
numeral. NumP is not projected in mass noun structures, and that is reflected
in their distribution (no pluralization, no direct combination with numerals).

(66)

```
                              DP
                 ╱────────────────────────╲
               D                          nP
               │                      ╱─────────╲
        ⎡ + atomic  ⎤               n            √
        ⎢           ⎥               │
        ⎣ − additive ⎦      ⎡ + collective ⎤
```

Finally, we should mention 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, namely, the broken plural shape and this agreed in the singular only, just like collective nouns do 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.

Next, we turn to the question as to why broken plurals, when interpreted as groups, correlate with feminine agreement on the verb (or the adjective). Feminine agreement appears not only on the verb when the verb is inflected singular, but also on adjectives and other such categories. In (67), the agreement on the quantifier *kolli* 'all' and the adverbial expression *mʕa bʕað-ha* 'together' are inflected in the feminine singular. This indicates that agreement is controlled by the subject noun, which is, as we argue in view of this and other evidence, syntactically feminine.

(67) *el wled   el kolli   xerj-et   mʕa bʕað-ha.*   [TA]
　　 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 agreement is continuous in the cases at hand.[13]

13　A reviewer asks whether one probe (say, an adjective) can agree syntactically while another (say, a verb) agrees semantically? The answer is no. At least, for Arabic. In that language, mixed agreement inside and outside DP is not possible. This is consistent with the behaviour of group nouns. In French, group nouns are attested (i), but there cannot be mismatches between verbs and adjectives in relation to the hybrid noun, as seen in (ii).

(i)　*Le jury, ils   prennent le   melon.*　　　　　　　　　　　　　[French]
　　 the jury they take   the melon
　　 'The jury, they are getting a big head.' (*Le Parisien*, quote from a candidate after a reality show, 2013, cited in Gardelle, 2019, p. 34)

(ii)　*Le nouveau/*nouveaux jury, ils   prennent le   melon.*　　　　[French]
　　 the new.SG/new.PL   jury they take   the melon
　　 'The new jury, they are getting a big head.'

On the other hand, hybrid nouns of the type found in Hebrew, e.g., *be'alim* 'owner', as discussed by Landau (2015), triggers either syntactic or semantic agreement within the DP. Although the Hebrew hybrid noun *be'alim* 'owner' is morphologically marked as masculine plural, it is fully compatible with both singular and plural referents, as shown in (iii) and (iv). (iii) and (iv) involve matching between the features on the verb and the adjective

(68) *ʕand-ek   meʃekel   nafsiy-ya.*   [Tunisian Arabic]
have-2SG problem.BPL personal-FEM.SG
'You have personal problems.'

(69) *xlaʕt   el beb beʃ taʕmel   el ʃweh*
force.PERF.2.SG the door will do.IMP.2.SG the scandal.BPL
heð-i   lkol?
this-FEM.SG all
'You forced the door to make all these scandals?'

Unlike sound plural suffixes, gender does not surface on broken plural forms (Caubet et al. 1989; Wright 1933:183). (70) illustrates that even when gender is visible on singular nouns such as *kalb*, 'dog', and *mrið*, 'sick', there exists only

---

whereas (v) does not: here, the verb is marked singular while the adjective is marked plural. (vi), where the features on the verb are plural but the features on the adjctive singular, is not possible:

(iii) *ha-beʼal-im   ha-kodem   maxar et   ha-makom lifney šana.* [Hebrew]
the-owner-PL the-previous.SG sold.3SG ACC the-place   before year
'The previous owner sold the place a year ago.'

(iv) *ha-beʼal-im   ha-kodm-im   maxru et   ha-makom lifney šana.*
the-owner-PL the-previous-PL sold.3PL ACC the-place   before year
'The previous owners sold the place a year ago.'

(v) *ha-beʼal-im   ha-xadavsim hexlit   al picul.*
the-owner-PL the-new.PL   decided.3SG on demerger
'The new owner decided on demerger.'

(vi) **ha-beʼal-im   ha-xadavs hexlitu   al picul.*
the-owner-PL the-new.SG decided.3PL on demerger
'The new owner(s) decided on demerger.'

According to Landau (2015), the patterns described above for Hebrew are made possible because hybrid nouns come with two types of features: morphologically-rooted (=CONCORD) features (hosted on the noun stem) and semantically-rooted (=INDEX) features (hosted on Num, a higher functional head). We adopt his discussion of Serbian/Croatian *deca* 'children' and propose that, like Serbian/Croatian *deca*, broken plurals in Arabic only make their CONCORD features available to attributive agreement. This follows from a selectional parameter: some nouns involve complete independence of the INDEX number from the CONCORD number (Hebrew *beʼalim*), standard nouns have a default CONCORD-INDEX matching constraint attached, and exceptional nouns of the *deca*-type, where the INDEX number is fixed.

Finally, (vi) is not possible because verbal agreement always originates from semantic features. It is unlike attributive adjectives that can pick up the features from either syntactic or semantic agreement. In (vi), the plural marking on the verb indicates that the INDEX value is plural. Therefore, if the agreement on the attributive adjective is singular, then it necessarily originates from the CONCORD value, and this is impossible, since *beʼalim* is morphologically plural.

one form for the broken plural, and, without a proper context, one could not tell which gender it refers to.

(70) a. *kalb*   *kalb-a*   *kleb*   [Tunisian Arabic]
     dog.MASC.SG dog-FEM.SG dog.BPL
     'dog (male), dogs (female), dogs'

   b. *mrið*   *mrið-a*   *morða*
     sick.MASC.SG sick-FEM.SG sick.BPL
     'sick, sick, sick'

Since the feminine is not morphologically marked on broken plurals, the only way to detect gender is through the agreement on the verb and other lexical items controlled by the noun. For example, the adjective *nafsiyya*, 'personal' in (68) is feminine singular, despite the fact that it modifies *meʃekel*, 'problems', a feminine broken plural. In (69), the broken plural noun *fweh*, 'scandals' is the antecedent to a feminine singular deictic pronoun *heði*. Deflected agreement, therefore, simply reflects the feminine gender that does not surface on the broken plural subject. This observation is not surprising, since based on the standard definition of gender by Hockett (1958), gender is reflected in the behaviour of associated words, rather than on the noun itself. This view is the predominant one in the literature and many authors adopted it in their analyses (see also Corbett, 1991; Aikhenvald, 1991; Kramer, 2009, 2015).

   As already hinted, we propose that the feminine in the cases at hand coorresponds to a morpheme [+FEM], exponed as *-a*, and that it denotes a group. That the group interpretation correlates with a feminine feature can be seen independently in contexts such as professional (71) or ethnic/regional groups (72). In each case, a gender shift from masculine to feminine yields a group interpretation.

(71) a. *ħajjem*   [Tunisian Arabic]
     hairdresser.MASC.SG
     'hairdresser'

   b. *ħajjem-a*
     hairdresser-FEM.SG
     'female hairdresser or a group of hairdressers'

(72) a. *bedwi*
         bedouin.MASC.SG
         'Bedouin'

      b. *bedwi-a*
         bedouin-FEM.SG
         'Female bedouin or a group of bedouins'

The feminine morpheme, phonologically spelled out as *-a*, appears on the noun. In this case, agreement on the verb is masculine (73). This is either because marking on the noun is sufficient for marking of the group denotation or because the agreement is automatically done semantically to avoid ambiguity, since feminine marking on the verb, adjective, etc. would express a different semantics, namely that a female bedouin travelled.

(73) *el   bedwi-a   sefr-u.*   [Tunisian Arabic]
      the bedouin-FEM.SG travelled-MASC.PL
      'The (group of) bedouins travelled.'

We can think of the feminine morpheme (exponent *-a*) as denoting a number rather than a gender feature, since it is correlated with the meaning of a group. There is evidence that the suffix *-a* was not originally associated with gender. In particular, there is an argument from reconstruction, that Proto-Semitic did not have gender (Hasselbach, 2014b) and that it developed from the *-a* suffix associated with nominalization. (74) summarizes what came first and next.

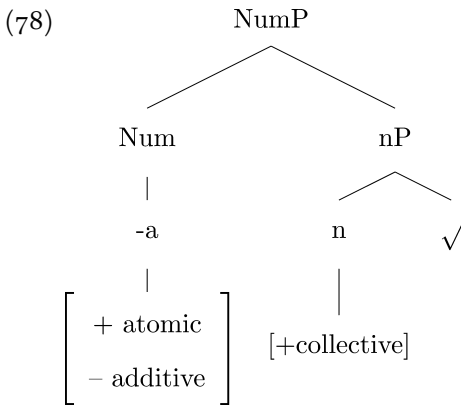(74) nominalization > singulative > group > gender

This diachronic process leads to a situation where the exponent *-a* is synchronically ambiguous in Arabic. It refers to different grammatical functions. As shown by (75), *-a* can, in TA, turn an adjective into a noun (there is additional stem internal change) along the lines of an operation like (76). This was the first use of the suffix *-a* (synchronically, nominalization commonly involves feminization, Kramer 2015, Alexiadou, 2004).

(75) *saʕeed saʕad-a*
      happy happiness
      'happy, happiness'

(76)   nP

```
        nP
       ╱⌒
      n    aP
      |    ⌒
     -a   a   √
```
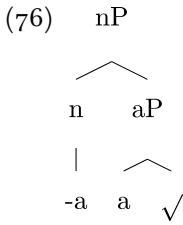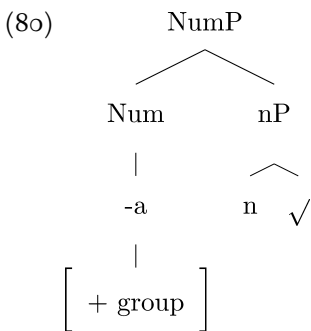
Next, we know from reconstruction that *-a* was used as a singulative marker, turning a collective noun into a singular. Consider (77) and (78).

(77) *beeð   beeð-a*
     egg.COLL egg-SING
     'eggs, an egg'

(78)

```
                    NumP
                  ╱⌒⌒⌒
              Num          nP
               |          ⌒
              -a        n    √
               |        |
    ⎡ + atomic  ⎤   [+collective]
    ⎣ − additive ⎦
```

Finally, *-a* was used to create groups from singular nouns, as in (79).

(79) *jazzar   jazzar-a*
     butcher.MASC.SG butcher-pl
     'butcher, butchers'

(80)

```
              NumP
             ╱⌒
        Num        nP
         |        ⌒
        -a       n   √
         |       |
   ⎡ + group ⎤
```

Broken plurals are very similar to (80) except that *-a* does not surface on broken plurals, but as we have seen, on adjectives, verbs, etc.

Finally, *-a* became a gender marker. We know from reconstruction studies that gender in Arabic first developed through agreement, most notably on adjectives, before appearing on substantives (Hasselbach, 2014b; Speiser, 1936).[14] A synchronic reflex of this is that, although adjectives will necessarily be marked for gender, it is not necessarily the case for substantives. This is illustrated in (81), where the feminine nouns *ʃams* 'sun' and *ʕin* 'eye' are not marked for gender but trigger feminine agreement on the adjectives.

(81) a. *ʃams qweyy-a*
      sun   strong-FEM.SG
      'a strong sun'

    b. *ʕin xaðr-a*
      eye green-FEM.SG
      'green eye'

Summarizing, we see from the above discussion that *-a* started out as a nominalizer, then was used as a number feature, and it is only later that it was used as a gender feature. While the exponent *-a* evolved diachronically, filling one function at a time, the synchronic picture is different. All the functions realized by the exponent *-a* are represented in the modern dialects. Generally speaking, all that the suffix *-a* is offering is marking a morphological contrast. In the non-collective class, it marks the feminine [+FEM] in singulars and group readings in the broken plural. In the collective class, it marks the singulative, as opposed to the collective.

Synchronically, we have a feature [+FEM] (a morpheme) and the features associated with [+FEM] are conditioned by the base of attachment. This can be justified by a weak allosemy scenario, where [+FEM], exponed by *-a*, is interpreted according to the following rules:

(82) LF instructions: semantic realizations of [+FEM]
    a. [+FEM] ↔ "singulative"/_____ $n_{[+\text{COLL}]}$
    b. [+FEM] ↔ "nominalizer"/_____ $n_{\text{INANIMATE}}$
    c. [+FEM] ↔ "nominalizer" and "female" /_____ n_{ANIMATE}

---

14   As is well known, "[g]enders are classes of nouns reflected in the behavior of associated words" (Hockett, 1958, p. 231).

In addition to the rules in (82), we add that the broken plurals are also endowed with a group feature, translated into a feminine feature [+FEM] and exponed by *-a*. All these synchronic parallel uses of the feminine feature [+FEM] reflect the different uses of the feminine throughout the historical changes. Although figuring out the meaning of the feminine feature [+FEM] and its exponent *-a* may seem like a daunting task for learners of the language, one must keep in mind that it simply depends on the basis to which it is attached, as suggested above.[15]

In section 4, we gave an analysis for the alternative ways broken plurals agree with the verb or adjective in Tunisian Arabic. We argued that broken plurals are hybrid nouns: either their syntactic features enter into an agree relation with the verb or adjective or their semantic features are accessed and the agreement is not syntactic, but semantic. For this to work in Tunisian Arabic, it was argued that gender features are visible at LF. We know this is possible independently (Hammerly, 2018) and this feature is therefore not an ad hoc feature of our analysis.

5 Conclusion

The aim of this paper was to account for an apparent agreement mismatch in Tunisian Arabic: the ϕ-features of broken plural subjects do not always seem to agree with the verb in gender and number. We argued that broken plurals in TA are hybrid nouns: either agreement is with the syntactic or the semantic features. We argued that broken plurals denote groups and that the marker *-a* is the spell out of a group feature. It was also shown that broken plurals have a very similar structure to that of collectives. Finally, it was argued that all gender features are visible at LF.

---

15 A reviewer suggested that the forms seen in feminine, group and singulative nouns are a case of metasyncretsim, since it holds across all agreement targets. We instead use the term weak conditioned allosemy, since it accounts for the fact that the meaning associated with the feature [+FEM] depends on certain characteristics of the base. In Distributed Morphology, as the reviewer correctly points out, if the data contains an exponent interpreted in different ways, that means it is an underspecified Vocabulary Item used to realize several different syntactic/semantic feature bundles, i.e., this is syncretism. Instead, we propose for the data at hand an allosemy rule. Allosemy rules, operate over feature bundles or abstract roots, not exponents like *-a*.

Acknowledgments

References

Acquaviva, P. (2008). *Lexical plurals: A morphosemantic approach*. Oxford: Oxford University Press.

Aikhenvald, A. (1991). *Classifiers: A typology of noun categorization devices*. Oxford: Oxford University Press.

Alexiadou, A. (2004). Inflection class, gender and DP-internal structure. In G. Müller, L. Gunkel, & G. Zifonun (Eds.), *Explorations in nominal inflection*. Berlin: Mouton. 21–50.

Arregi, K. and A. Nevins (2012). *Morphotactics: Basque auxiliaries and the structure of spellout*. Dordrecht: Springer.

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.

Blanc, H. (1970). Dual and pseudo-dual in the Arabic dialects. *Language 46*, 42–57.

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.

Boris, G. (1945–1948). Sur l'emploi des pluriels féminins dans un parler arabe moderne. *Comptes Rendus du Groupe Linguistique d'Études Chamito-sémitiques (GLECS)*, 21–23.

Brustad, K. (2000). *The syntax of spoken Arabic*. Georgetown University Press.

Caubet, D., M.-C. Simeone-Senelle, and M. Vanhove (1989). Genre et accord dans quelques dialectes arabes. LINX 21: Genre et langage. Actes du colloque tenu à Paris X—Nanterre, les 14-15-16 décembre 1988.

Comrie, B. (1975). Polite plurals and predicate agreement. *Language 51*(4), 406–418.

Corbett, G. (1991). *Gender*. Cambridge University Press.

Corbett, G. (2000). *Number*. Cambridge University Press.

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.

Croft, W. (2003). *Typology and universals*. Cambridge University Press.

Dali, M. and E. Mathieu (2020). *A theory of distributed number*. John Benjamins, To appear.

D'Anna, L. (2017). Agreement in plural controllers in Fezzānī Arabic. *Folia Orientala LIV*, 101–122.

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 Vries, H. (2013). Distributivity and agreement: New evidence for groups as sets. In M. Aloni, M. Franke, and F. Roelofsen (Eds.), *Proceedings of the 19th Amsterdam Colloquium*, pp. 241–248.

de Vries, H. (2015). *Shifting sets, hidden atoms: The semantics of distributivity, plurality and animacy*. Doctoral dissertation, Utrecht University.

den Dikken, M. (2001). Pluringulars, pronouns and quirky agreement. *The Linguistic Review 18*, 19–41.

Doron, E. and A. Müller (2013). The cognitive basis of the mass-count distinction: Evidence from bare nouns. In C. Hofherr and A. Zribi-Hertz (Eds.), *Crosslinguistic studies on noun phrase structure and reference*, pp. 73–101. Brill.

Enger, H. (2004). Scandinavian pancake sentences as semantic agreement. *Nordic Journal of Linguistics 3*, 1–54.

Enger, H. (2013). Scandinavian pancake sentences revisited. *Nordic Journal of Linguistics 36*, 275–301.

Fassi Fehri, A. (1988). Agreement in Arabic, binding and coherence. In M. Barlow and C. Ferguson (Eds.), *Agreement in natural language: Approaches, theories, description*, pp. 107–158. Stanford, CA: CSLI.

Fassi Fehri, A. (2003). 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.

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.

Gardelle, L. (2019). *Semantic Plurality*. John Benjamins Publishing Company.

Grimm, S. (2012). Inverse number marking and individuation in Dagaare. In D. Massam (Ed.), *Count and mass across Languages*, pp. 75–98. Oxford: Oxford University Press.

Grimshaw, J. (2005). Extended projection. In J. Grimshaw (Ed.), *Words and structure*, pp. 1–74. CSLI Publications.

Haelewyck, J.-C. (2016). *Grammaire comparée des langues sémitiques*. Bruxelles, éditions Safran.

Hammerly, C. (2018). Limiting gender. In E. Mathieu, M. Dali, 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. (2011). Valence and atomic number. *Linguistic Inquiry 42*, 561–594.

Harbour, D. (2014). Paucity, abundance, and the theory of number. *Language 90*, 185–229.

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.

Hockett, C.F. (1958). *A course in modern linguistics*. New York: MacMillan.

Holes, C. (1990). *Gulf Arabic*. London: Routledge.

Johnson, C.A. and B.D. Joseph (2014). Morphology and syntax and semantics and pragmatics … deconstructing "semantic agreement". *Lingvisticæ Investigationes 37*, 306–321.

Kihm, A. (2003). Les pluriels internes de l'arabe: Système et conséquences pour l'architecture de la grammaire. *Recherche Linguistiques de Vincennes 32*, 109–156.

Kramer, R. (2009). *Definite markers, phi-features, and agreement: A morphosyntactic investigation of the Amharic DP*. Ph.D. thesis, UC Santa Cruz.

Kramer, R. (2015). *The morphosyntax of gender*. Oxford: 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. syntax. In F. Cooley (Ed.), *The Proceedings of the 17th Texas Linguistic Society*, pp. 39–53.

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 (2015). On plurals, noun phrase and num(ber) in Moroccan Arabic and Djibouti Somali. In S. Bendjaballah, N. Faust, M. Lahrouchi, and N. Lampitelli (Eds.), *The form of structure, the structure of form: Essays in honor of Jean Lowenstamm*, pp. 303–314. Amsterdam: John Benjamins.

Landau, I. (2015). Dp-internal semantic agreement: A configurational analysis. *Natural Language and Linguistic Theory 34*, 975–1020.

Lecarme, J. (2002). Gender "gender" polarity: Theoretical aspects of Somali nominal morphology. In P. Boucher and M. Plénat (Eds.), *Many morphologies*, pp. 109–141.

Levy, M.M. (1971). *The plural of the noun in Modern Standard Arabic*. Ph.D. thesis, uitgever niet vastgesteld.

Lipiński, E. (2001). *Semitic languages outline of a comparative grammar. Second edition*. Leuven, Peeters.

Lowenstamm, J. (2008). On little n, √, and types of nouns. In J. Hartmann, V. Hegedus, and H. van Riemsijk (Eds.), *Sounds of silence: Empty elements in syntax and phonology*, pp. 105–144. Amsterdam: Elsevier.

Martí, L. (2018). Inclusive plurals and the theory of number. *Linguistic Inquiry*.

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. (2014). Many a plural. In A. Aguilar-Guevara, B.L. Bruyn, and J. Zwarts (Eds.), *Weak referentiality*, pp. 157–181. Amsterdam: John Benjamins.

Matushansky, O. (2013). Gender confusion. In L.L.-S. Cheng and N. Corvert (Eds.), *Diagnosing syntax*, pp. 271–294. Oxford: Oxford University Press.

McCarthy, J. and A. Prince (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory 8*, 209–283.

Mohammad, M.A. (1990). The problem of subject-verb agreement in Arabic: Towards a solution. *Perspectives on Arabic linguistics 1*, 95–125.

Murtonen, A. (1964). *Broken plurals, origin and development of the system*. Brill Archive.

Noyer, R. (1992). *Features, positions and affixes in autonomous morphological structure*. Ph.D. thesis, MIT.

Ojeda, A. (1992). The semantics of number in Arabic. *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory 2*, 303–326.

Ouwayda, S. (2014). *Where number lies: Plural marking, numerals, and the collective–distributive distinction*. Doctoral dissertation, University of Southern California.

Percus, O. (2011). Gender features and interpretation: A case study. *Morphology 21 (2)*, 167–196.

Preminger, O. (2014). *Agreement and its failures*. Cambridge, MA: MIT Press.

Ratcliffe, R. (1998). *The 'Broken' plural problem in Arabic and comparative Semitic*. Amsterdam: John Benjamins.

Smith, P. (2015). *Feature mismatches: Consequences for syntax, morphology and semantics*. Doctoral dissertation, University of Connecticut.

Smith, P.W. (2017). The syntax of semantic agreement in English. *Journal of Linguistics 53*(4), 823–863.

Speiser, E.A. (1936). Studies in semitic formatives. *Journal of the American Oriental Society 56*, 22–46.

Wechsler, S. and L. Zlatić (2003). *The many faces of agreement*. Stanford: CLSI.

Wright, W. (1933). *A Grammar of the Arabic Language (Volume I)*. Cambridge, MA: MIT Press.

Zabbal, Y. (2002). *The semantics of number in the Arabic noun phrase*. MA dissertation, University of Calgary.