Agreement Patterns in Khuzestani Arabic

By SIMONE BETTEGA (Università degli Studi di Torino) and BETTINA LEITNER (University of Vienna)

Abstract
In the course of this paper we analyze the functioning of agreement in the variety of Arabic currently spoken in the Iranian province of Khuzestan. Since Khuzestani Arabic distinguishes gender in the plural forms of the adjective, verb and pronoun, we adopt the theoretical framework first outlined in Bettega (2019) for the analysis of agreement patterns in gender-distinguishing varieties of Arabic. As in other Arabic dialects that preserve a M/F distinction in the plural, in Khuzestani Arabic non-human referents that trigger masculine agreement in the singular attract feminine agreement in the plural. This behavior can be accounted for by postulating the existence of three distinct agreement classes (or controller genders). The occurrence of singular agreement with plural controllers, on the other hand, should be understood in terms of variation in the level of individuation of the referent. In the article, we analyze the effects that several different factors have on agreement, including: the semantics of the controller, quantification, word order, distance between target and controller, and target type. Speaker’s age and the possible interference of typical MSA agreement patterns are also taken into account.

Keywords: agreement, Khuzestani Arabic, gender, plurality, morphosyntax, dialectology

Addresses for correspondence: simone.bettega@unito.it; bettina.leitner@univie.ac.at

1. Introduction
In this paper, we aim at offering a description of gender and number agreement patterns in Khuzestani Arabic (KhA). The present work appears in the wake of a number of studies, all published in recent years, that have brought attention back to the relatively neglected topic of agreement in Arabic. It is probably the case that agreement came to be considered a somehow secondary object of investigation because its general behavior in Arabic was – in theory – already well described and understood. Agreement, in particular, has often been used as a diagnostic tool to tell apart (and lump together) the typologically “innovative” spoken dialects from the putatively “conservative” standard version of the language (see for instance Fischer & Jastrow 1980: 47, 96, Versteegh 1984: 103-5, and more recently Corriente 2008: 20-21). However, in the course of the last decades, several in-depth studies on the topic have contributed in unearthing an unexpected degree of complexity, thus warranting a re-examination of the whole question.

* Simone Bettega is the author of §1 and §4 of the present paper, Bettina Leitner is the author of §2; §3 has been written together.
On the one hand, studies focused on the written version of the language (Belnap & Shabaneh 1992, Belnap & Gee 1994, D’Anna 2017b, D’Anna, forth.) have brought into focus its lack of diachronic homogeneity. In particular, the earliest attestations of Arabic (such as the pre-Islamic poetic corpus and the Quran) show agreement patterns that not only radically differ from those found in contemporary Modern Standard Arabic (MSA)\(^1\), but also bear striking similarities to the ones encountered in several modern dialects (see below).

On the other hand, the pioneering works of Belnap (1991, 1993) and Brustad (2000) employed new methodologies to tackle the issue of agreement in Spoken Arabic (SA), showing how, in this respect, the oral varieties are characterized by a high degree of inter- and intra-dialectal variation. As far as the latter is concerned, Belnap and Brustad both concluded that speaker choices in matters of agreement are determined by a number of complex and interrelated (but typologically common and at least partially predictable) factors (see below and §3).

Belnap’s and Brustad’s new approaches renewed the interest in agreement in SA, and a number of studies have recently appeared that adopt their methodologies, focusing on specific dialectal areas within the Arab world. Belnap’s studies on Cairene, in particular, and Brustad’s comparative analysis of four different dialects (Moroccan, Egyptian, Syrian and Kuwaiti), were all focused on dialects that have lost gender distinction in the plural forms of the adjective, verb and pronoun. The same holds true for the more recent works of Procházka and Gabsi (2016) on urban Tunisian and Holes (2016: 326-354) on Bahraini.

Herin and Al-Wer (2013), Ritt-Benmimoun (2016), Bettega (2017, 2018) and D’Anna (2017a), on the contrary, all concentrate on varieties of SA where this distinction has been maintained. KhA, which is the object of the present article, belongs to this second group.

Bettega (2019) argues that the divide between gender-distinguishing and non-distinguishing varieties\(^2\) is of fundamental importance, because the former show

---

\(^1\) This consideration, though often forgotten in subsequent works, is anything but new. Beeston (1975: 65-6) was already remarking how: “the use of the feminine singular concord with “irrational” substantives is a neologism in Arabic which only gradually won its way to becoming the norm”. Even Wright, in his authoritative grammar, noted that “when the subject in the plural denotes irrational or inanimate objects, the plur. fem. of the verb is preferred in classical Arabic, if their number does not exceed ten” (Wright 1896: 293).

\(^2\) For the sake of practicality, through the course of this article we will employ the labels “gender-distinguishing” and “non-distinguishing” to refer to dialects of Arabic that have, respectively, retained or lost gender distinction in the plural. This terminology obviously implies a certain degree of approximation, since all dialects of Arabic preserve gender distinction in the singular (at least in the third person).
a kind of agreement behavior entirely different from that of latter. In particular, Bettega puts forward the hypothesis that three separate agreement classes exist in gender-distinguishing varieties of SA, in spite of the fact that only two sets of morphological markers exist in Arabic to mark gender and number. The first of these classes only includes nouns with biologically masculine human referents: these attract masculine agreement both when singular and when plural. The second class contains all nouns referring to biologically feminine human beings, along with many nonhuman and inanimate ones. The vast majority of its members are overtly marked for feminine in the singular, and they attract feminine agreement in the singular and feminine agreement in the plural. The third class is anomalous, in that it has no dedicated set of morphological markers: it contains all nonhuman nouns that are not part of Class II, and these trigger masculine agreement when singular but feminine agreement when plural (cf. Wright quoted in fn. 1 above).

It is indeed beyond doubt that, in all gender-distinguishing dialects for which we have reliable data, whenever a plural nonhuman controller triggers plural agreement, it is always feminine agreement, and never masculine\(^3\). The same seems to hold true for KhA: this will be discussed in more depth in §3, where we will offer a description of agreement patterns in this dialect, and check the tenability of Bettega’s hypothesis in light of the data presented.

Apart from plural agreement, another agreement option seems to be available for plural controllers in almost every contemporary dialect of Arabic (gender-distinguishing or not): feminine singular agreement. Several studies on the topic all concur in indicating a referent’s level of individuation as the main cause of vacillation between plural (be it M. or F.) and F.SG. agreement (see Belnap 1991\(^4\), Brustad 2000, Holes 2016, Ritt-Benmimoun 2016 and Bettega 2017, to cite but a few). Since individuation will figure importantly in the rest of our discussion, it is worth briefly expanding on the concept here.

\(^3\) Throughout the Arabic-speaking world, many gender-distinguishing varieties of Arabic are currently in the process of losing this feature. This might be due to contact with other, locally prestigious non-distinguishing varieties, as well as other sociolinguistic factors and/or internal development. In these cases, agreement patterns in the plural may appear somehow erratic. Several independent sources, however (Ritt-Benmimoun 2016: 283, Hanitsch 2011: 148 and Ingham 1994: 64-5), confirm that, whenever this happens, feminine plural agreement is more often replaced by masculine plural agreement in the case of human controllers. Nonhuman controllers, on the contrary, appear to be more resistant to this process of syntactical reduction (though they are also affected to an extent).

\(^4\) Belnap did not actually employ the label “individuation”, which was first used in reference to Arabic agreement patterns by Brustad. However, many of the factors he indicated as capable of affecting agreement are actually the same ones that Brustad will later rubricate under this heading.
Individuation (and salience) are umbrella-terms used by typologists to refer to a number of distinct but interrelated factors which appear to have a bearing on several different linguistic phenomena (grammatical agreement being one of those). Comrie (1989: 19) writes that “salience relates to the way in which certain actants present in a situation are seized on by humans as foci of attention, only subsequently attention being paid to less salient, less individuated objects […]. Salience is not treated as a primitive in itself, but rather as the result of the interaction of a number of factors, such as animacy in the strict sense, definiteness, singularity, concreteness, assignability of a proper name”. Khan (1984: 469-470) was the first to apply the concept to the field of Semitic linguistics, adding an important remark to Comrie’s list, namely the fact that specific entities are more individuated than generic nominals referring to a whole class of entities. He also noted how textual salience (i.e. whether or not the referent has a prominent role within the text) can affect agreement. In her book, Brustad (2000: 22-5) uses individuation to explain the apparently erratic agreement patterns of plural nouns in SA. She adapts Khan’s list of factors by making a number of important additions. One is physical prominence, which parallels Khan’s idea of textual prominence; another is quantification, which can be used to contrast “collectivity or generality” with “heterogeneity and particularity” (numerals below ten, in particular, apparently tend to emphasize the latter: here as well, see fn. 1 above).

Other explanations have been put forward for the frequent use of feminine singular agreement with plural referents in SA. Owens and Bani-Yasin (1987), for instance, ascribe the presence of this particular pattern in the rural Jordanian dialect they analyzed to the influence of MSA. The MSA prescriptive rule that mandatorily requires F.SG. agreement with nonhuman plural referents, thus, may be “creeping in” the spoken language, along with the loanwords which more and more frequently are being borrowed from MSA. This, however, does not explain why even some human controllers do sometimes attract F.SG. agreement, or why, even in the case of nonhuman ones, it is not always loanwords from MSA that trigger this specific pattern, but also some “purely” dialectal items (see Bettega 2018 for a discussion).

Herin and Al-Wer (2013) do similarly maintain that F.SG. agreement with plural referents is an innovation caused by contact in the Jordanian dialect of Salt. However, they indicate other locally prestigious (but non-distinguishing) dialects as the source of this peculiar pattern (the “original” dialectal rule being F.PL. agreement with nonhuman controllers, and plural agreement, either M. or F. depending on their biological sex, for human ones).

In §3 we will come back to this hypothesis, and in general to the topic of F.SG. agreement with plural controllers, discussing how the KhA data presented can help us to better understand this phenomenon (also in terms of its diachronic development). Before moving to data analysis, however, we will offer a brief sketch of the sociolinguistic profile of KhA.
2. Historical Development of Arabic in Khuzestan

The Iranian province of Khuzestan is situated in the south-west of the country at the border to Iraq. First settlement of Arab tribes in this region preceded the Arab destruction of the Sasanian empire with the rise of Islam. For example, the Banū Tamīm, reportedly had settled in Khuzestan prior to the arrival of the Arab Muslim armies (Daniel 1986: 211; cf. Zarrinkūb 1975: 27). In the centuries after the initial Islamic victories in the region, and especially after the capture of Šūštar 641-2, Khuzestan fell city by city to the Arabs (Zarrinkūb 1975: 15; Oppenheim 1967: 6) who established garrisons in the newly conquered region (Zarrinkūb 1975: 27). In the subsequent centuries, further nomadic Arab tribes like the Ḥanīfa, Tamīm, ʿ Abd-al-Qays, Banū Lām, Kašb, Āl Kaḥr and Āl Xamīs crossed the Persian Gulf or moved on from southern Iraq into Khuzestan (Oberling 1986: 215-16; Oppenheim 1967: 6, 9-11). As a consequence of this influx of Arabic speakers, the region was called ʿArabistān since the reign of Shāh Šāh Abbās the Great (1588-1629) until 1925 (cf. Gazsi 2011: 1020).

From the early 20th century on, Khuzestan has been of particular interest to the Persian government due to its abundant oil resources and has attracted international, especially British, interest. Many Arabs and Persians living in Khuzestan work in the sugar cane or oil industries (de Planhol 1986: 55-56), but agriculture (e.g. palm groves and date production, especially around Fālāḥiyā, and buffalo breeding by the Miḍān in the marshes close to the Iraqi border) is also still a major source of income in Khuzestan.

Though there is no exact number of the current number of Arabic speakers in Iran, it is estimated that roughly 4% of Iran’s citizens, 3 million people, are Arabs. The majority of these, around 2.5 million, live in the province Khuzestan (Matras & Shabibi 2007: 137; Gazsi 2011: 1020). Many villages and towns in Khuzestan are populated mainly by Arabs and some of them are inhabited by Arabs only.

In the past decades, there has been considerable movement of Khuzestan’s inhabitants as a consequence of the Iran-Iraq war (1980-88), or due to socio-economic reasons. People migrated within the province (e.g. from villages to towns) as well as to other countries such as Iraq or Kuwait.

---

5 Cf. Oberling (1986: 218), Oppenheim (1967: 3-97) and Layard (1846: 37) for an overview of the Arab tribes in Khuzestan.

2.1 Typology of Khuzestani Arabic (KhA)
KhA belongs to the Bedouin-type South-Mesopotamian _galat_-dialects. Therefore, it shows great similarity to (southern) Iraqi dialects of the Bedouin type, e.g. Basra Arabic and Muslim Baghdadi Arabic, as well as to Gulf dialects like Bedouin (i.e. Sunni) Bahraini Arabic.

KhA can be considered a “peripheral” dialect of Arabic because it is spoken in a country where Arabic is not the language of the majority population and is not used in education and administration. Consequently, the influence of Modern Standard Arabic in education is limited to few and special contexts, e.g. the teaching of Classical Arabic in Quran-classes.

However, due to its long, geographically open border with Iraq, Khuzestan is not totally isolated from the Arabic-speaking world and many Arabs from Khuzestan regularly visit Iraqi towns, e.g. Kerbala, for religious purpose like visits of the holy shrine of Imam Husain.

Via satellite TV, the Arabic community of Khuzestan has had access to Arabic news, soaps, etc. since around 2000. The broadened access – via television and of course also via the internet – to the rest of the Arabic-speaking world might have a growing impact on the language usage of KhA-speakers and could lead to the abandonment of archaic features such as the retention of F.PL. forms.

The knowledge about oral tradition and the ability to tell folk stories by heart is nowadays mostly limited to the eldest female members of the language community and will probably disappear soon as its popularity has been declining in the recent decades. Many KhA speakers themselves attribute this phenomenon to the popularity of (satellite) television, which has grown to be a major source of entertainment.

Poetry in the vernacular language, however, is still very popular among KhA-speakers and its preservation is fostered by poetry competitions as well as popular publications on vernacular KhA poetry.

2.2 Language Policy and Bilingualism
The Western Iranian language Persian is the only official language in Iran and the only language used in education. It is sociolinguistically and culturally dominant (cf. Hayati and Mashhadi 2010: 27), especially in the domains of business and administration. Consequently, the Arabic language has no institutional support in Iran outside religious instruction.

---

7 There is as yet no comprehensive grammar of KhA. For now, the main source of information on KhA is the collection of data made in the 1960s by the arabist and linguist Bruce Ingham (2007; 1976; 1973). The article by Yaron Matras & Maryam Shahibi, “Grammatical borrowing in Khuzistani Arabic” (2007), is based on Shahibi’s unpublished dissertation Contact-induced grammatical changes in Khuzesti Arabic (2006).
The majority of KhA speakers in Khuzestan and all speakers recorded for collecting the data used in the present article are cognitively dominant in KhA and thus have KhA as their L1.8

Some of the recorded informants belonging to the older generation can be considered monolinguals with almost no knowledge of Persian. However, as a consequence of the Persian monolingual education system, most Arabic speakers become bilinguals and acquire a good level of spoken as well as written Persian. It has to be kept in mind that the degree of bilingualism or proficiency of Persian among KhA speakers can vary to a certain degree (cf. Matras & Shabibi 2007: 147) depending on factors such as level of education, occupation, age, and gender. For example, the limited access to education and jobs among the older generation and women implies less exposure to language contact situations with Persian and a lower degree of bilingualism.

In most cases, KhA remains the dominant language at home, within the family and among community members whereas at work or in school the dominant language is Persian.

Persian enjoys high prestige in the Iranian society whereas KhA has very low prestige and is not associated with the highly prestigious Arabic of the Quran, which is taught at schools. Quite often, Khuzestan Arabic speakers themselves also perceive KhA as less prestigious than Classical Arabic.

Persian has had a strong influence on KhA lexicon as well as on its grammar.9 This can be attributed mainly to the long history of Arabic-Persian language contact in the area and to the dominance of Persian in the modern Iranian educational system, administration and work environment, which in turn has led to a widely bilingual society in Khuzestan.

3. Data analysis
In this paragraph, we will present the results of the analysis we ran on a corpus of Khuzestani texts. The texts were recorded by Bettina Leitner during a first fieldwork period in Khuzestan in 2016, followed by a second period spent among the Khuzestani community in Kuwait in 2018. Several informants residing in Vienna were consulted as well during the same period (2016-18). The entire transcribed corpus consists of about 28,000 words and is composed of 58 texts of different genres (fairy tales, everyday conversation/dialogues, childhood memories, traditions and customs, recipes), gathered with the help of 34 different speakers (18 males and 16 females) aged 27-ca.75 years.

8 Some Arabic families in Khuzestan raise their children with Persian hoping they would enhance in this way their children’s career prospects.

9 See Leitner (forthc.) for examples of Persian influence on Khuzestani phonology, morphosyntax (replication of Persian phrasal verb constructions), syntax (word order), and lexicon.
All targets that appeared in the corpus agreeing with a plural or collective controller were extracted and recorded in an Excel spreadsheet. This yielded a total of 270 controllers with 521 corresponding targets. These were later coded for a number of relevant morphological, syntactical and semantic traits, and a separate statistical analysis was run for each one of these, in order to assess their role in determining agreement outcomes. All of these factors will be dealt with in detail in the following paragraphs, starting with controller-related factors and then moving on to target-related and speaker-related ones.

### 3.1 Controller-related factors: humanness, animacy and collectiveness

In his study on agreement in Cairene Arabic, Belnap (1991: 61-8) found humanness (or animacy) to be the one factor having the strongest influence on agreement outcomes\(^\text{10}\). All subsequent studies on the same topic in different varieties of Arabic basically confirmed this. We will here analyze humanness/animacy together with collectiveness for reasons that will be presently made clear.

Of the 270 controllers that constitute our corpus, 98 denote human referents, 41 denote animals, and 134 inanimate things/objects. 198, 73 and 250 targets depend on these three controller types, respectively. As shown in the tables below, different controller types tend to attract different types of agreement. In particular, controllers whose level of individuation is intrinsically low tend to attract F.SG. agreement more often than other controllers (the relation between F.SG. agreement and lack of individuation has been discussed in §1). Collective nouns are inherently characterized by low individuation, so that it would make little sense to analyze them together with proper plurals. This is why, in the following tables, collectives and non-collectives have been kept distinct.\(^\text{11}\)

\(^{10}\) Belnap analyzed the morphological status of the controller (i.e. its being a suffixal or apophonic plural) together with animacy. Occurrences of suffixal “sound” plurals were rare in our corpus compared to apophonic ones, so that statistical analysis is not particularly meaningful. In general, though, and unsurprisingly, it would seem that “broken” plurals are somehow more likely to attract F.SG. agreement than sound ones (in the case of human controllers, only 1 out 27 targets depending on sound plurals agreed in the F.SG., while 5 out of 77 of those depending on broken ones did. In the case of nonhuman controllers, the ratio was 7 out of 40 for targets depending on sound plurals against 60 out of 175 for broken ones).

\(^{11}\) For the sake of the present analysis we have adopted a working definition of “collective” based on mixed semantic and morphological criteria. We will consider collectives those nouns which are semantically referred to a group or plurality of some sort, and that, however, are not morphologically plurals (either because an actual plural, and sometimes a singulative, can be derived from them, or because they exist in a single invariable form that cannot be inflected for plural or singular or singulative). Also, collectives cannot be accompanied by a numerical quantifier.
### Table 1A – Targets depending on human controllers (collectives)

<table>
<thead>
<tr>
<th>Agreement type</th>
<th>M.SG.</th>
<th>F.SG.</th>
<th>M.PL.</th>
<th>F.PL.</th>
<th>BROKEN PL.</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of targets</td>
<td>8</td>
<td>23</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>Percentage</td>
<td>8.7%</td>
<td>25%</td>
<td>66.3%</td>
<td>0%</td>
<td>0%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 1B – Targets depending on human controllers (non-collectives)

<table>
<thead>
<tr>
<th>Agreement type</th>
<th>M.SG.</th>
<th>F.SG.</th>
<th>M.PL.</th>
<th>F.PL.</th>
<th>BROKEN PL.</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of targets</td>
<td>0</td>
<td>2</td>
<td>67</td>
<td>35</td>
<td>3</td>
<td>106</td>
</tr>
<tr>
<td>Percentage</td>
<td>0%</td>
<td>1.9%</td>
<td>63.2%</td>
<td>33%</td>
<td>2.8%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2A – Targets depending on animate nonhuman controllers (collectives)

<table>
<thead>
<tr>
<th>Agreement type</th>
<th>M.SG.</th>
<th>F.SG.</th>
<th>M.PL.</th>
<th>F.PL.</th>
<th>BROKEN PL.</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of targets</td>
<td>13</td>
<td>16</td>
<td>1</td>
<td>14</td>
<td>2</td>
<td>46</td>
</tr>
<tr>
<td>Percentage</td>
<td>28.3%</td>
<td>34.8%</td>
<td>2.2%</td>
<td>30.4%</td>
<td>4.3%</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2B – Targets depending on animate nonhuman controllers (non-collectives)

<table>
<thead>
<tr>
<th>Agreement type</th>
<th>M.SG.</th>
<th>F.SG.</th>
<th>M.PL.</th>
<th>F.PL.</th>
<th>BROKEN PL.</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of targets</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>19</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>Percentage</td>
<td>3.7%</td>
<td>14.8%</td>
<td>7.4%</td>
<td>70.4%</td>
<td>3.7%</td>
<td></td>
</tr>
</tbody>
</table>

---

12 Obviously, only adjectives can show broken plural agreement, so that they are more restricted in their distribution than the other agreement types. As a consequence of this, the figures shown in this column are only partially comparable with those appearing in columns 1-4, since these contain data relative to verbs and pronouns as well. A correlation has been repeatedly shown to exist between broken agreement (in adjectives) and F. SG. agreement (in other target types and/or other adjectives). This has been known since the earliest grammatical descriptions of the Arabic language (see for instance al-Mubarrad’s *Muqtadab*). Broken agreement would thus appear to have a more neutral status regarding the M. and F. distinction. Though the percentages of broken agreement could have perhaps been merged with those relative to F.SG. agreement, we decided to keep them separate in the tables.
Table 3A – Targets depending on inanimate controllers (collectives)

<table>
<thead>
<tr>
<th>Agreement type</th>
<th>M.SG.</th>
<th>F.SG.</th>
<th>M.PL.</th>
<th>F.PL.</th>
<th>BROKEN PL.</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of targets</td>
<td>19</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>Percentage</td>
<td>70.4%</td>
<td>22.2%</td>
<td>0%</td>
<td>7.4%</td>
<td>0%</td>
<td></td>
</tr>
</tbody>
</table>

Table 3B – Targets depending on inanimate controllers (non-collectives)

<table>
<thead>
<tr>
<th>Agreement type</th>
<th>M.SG.</th>
<th>F.SG.</th>
<th>M.PL.</th>
<th>F.PL.</th>
<th>BROKEN PL.</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of targets</td>
<td>13</td>
<td>69</td>
<td>3</td>
<td>134</td>
<td>4</td>
<td>223</td>
</tr>
<tr>
<td>Percentage</td>
<td>5.8%</td>
<td>30.9%</td>
<td>1.3%</td>
<td>60%</td>
<td>1.8%</td>
<td></td>
</tr>
</tbody>
</table>

These figures obviously require some explanation. The first thing that can be noted is that M.SG. agreement is not really an agreement option for proper plurals, while it is fairly common with collectives (in particular, out of 53 occurrences of M.SG. targets in the corpus, 40 depend on collective controllers). This is because, as we have seen (fn. 11), many collective nouns are morphologically akin to masculine singular ones. This means that KhA speaker have at least three different agreement options for collective nouns, that is, M.SG., F.SG. and plural (plural can then be masculine or feminine, depending on the referent). While oscillation between F.SG. and plural agreement can be explained in terms of variation in the degree of individuation, it is less clear what prompts speakers to opt for M.SG. agreement with collective nouns. It would seem that individuation plays a role here as well 13, since the increase in frequency is monotonic, starting low with human collectives (intrinsically more individuated) and growing more common with animal collectives and then inanimate ones 14.

---

13 This idea can already be found in Herin and Al-Wer (2013: 67), though the only element they report as capable of triggering three different types of agreement in the Jordanian dialect of Salt is nās, “people”. They note how only indefinite occurrences of nās can actually trigger M.SG. agreement. In our KhA data, the only occurrence of nās that attracts this type of agreement is indeed indefinite.

14 A few cases (13) of M.SG. agreement with non-collective controllers occur as well (though never with human ones). Some of these are due to the effects that word order has on agreement and will be discussed in §3.5, while others are instances of grammaticalized or grammaticalizing particles, that no longer show the ability to inflect, or are currently in the process of losing it (such as māl, “property”, often used as a preposition meaning “of, belonging to”). Others still are occasional “slips of the tongue” on the part of the speaker, who immediately corrects him/herself. Seven M.SG. targets appear in the data that depend on non-collective inanimate controllers and to which the above explanations cannot be applied. They all depend on the two
Considerations similar to those made for M.SG. agreement hold for F.SG. agreement. This type of agreement obviously obtains when the referent’s level of individuation is low, and again, increase in frequency is monotonic with non-collective controllers, starting low with humans and then growing more common with animal and inanimate controllers. Although more common with collective controllers (examples 1 to 3 below), F.SG. agreement can occur with all kinds of controllers, including human plurals (examples 4 and 5):

1) Ahwaz (Khuzestan), female, 30 years (Leitner, own data)

\[ \text{ən-nās gabul ťāstafād-ha.} \]

DEF-peoples before use.PRS.3F.SG-PRON.3F.SG

“In former times, the people used it.”

2) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

\[ \text{ahl əl-halāl ťāfārəf əl-ḥalāl-ha.} \]

family DEF-livestock know.PRS.3F.SG livestock-PRON.3F.SG

“A family who owns livestock knows its livestock.”

3) Ahwaz (Khuzestan), female, 50 years (Leitner, own data)

\[ \text{gabul məmša l-wādim ətrūh l-əl-ḥaġğ.} \]

before walk.PTCP.SG DEF-peoples go.PRS.3F.SG to-DEFPRES.3F.SG

“In former times people (used to) walk when they went on ḥaġğ.”

4) Ahwaz (Khuzestan), female, 50 years (Leitner, own data)

\[ \text{hāy əl-frūx əll-ha.} \]

DEM.F.SG DEF-children all-PRON.3F.SG

“These children, all of them.”

5) Hamidiyya (Khuzestan), female, 60 years (Leitner, own data)

\[ \text{ṣārat əd-hum əfrēx-āt.} \]

become.PRF.3F.SG at-PRON.3M.PL child.DIM-PL

“They got children.”

It is worth noting that the occurrence of F.SG. agreement with human plurals or collectives is in open violation of the prescriptive rules of MSA\textsuperscript{15}, and thus invalidates the theory (discussed in §1) that the use of F.SG. agreement with plural plurals malābis and əhdūm, both meaning “clothes”, and mostly consist of color adjectives. For an example of a collective controller triggering three different types of agreement, see example (22) below.

\textsuperscript{15} In MSA as well some human collectives can actually trigger F.SG. agreement (nās being the most common). This phenomenon, however, is much more limited in MSA than in the spoken dialects.
controllers in SA is the result of contact with the standard variety (on this point see, however, §3.6 below).

Another interesting remark concerning F.SG. agreement is that its increase in frequency is not monotonic when collective controllers are taken into account, since it appears to be more common with collectives referring to animals than with any other type of controller. This is unsurprising, if one considers that animals in languages throughout the world often come conveniently organized into herds, flocks, packs and the like, being perceived as compact groups or masses with no internal division or specification. Examples (6) and (7) are instances of animal collectives triggering F.SG. agreement:

6) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

duwāb kaṭīra
water_buffalos many.F.SG

“He many water buffalos.”

7) Ahwaz (Khuzestan), female, 50 years (Leitner, own data)

ləga əl-ɡanam kəll-ha mkəssara.
find.PRF.3M SG DEF-sheep all-PRON.3F SG broken.PTCP.F.SG

“He found the sheep, all of them broken-footed.”

Moving on to occurrences of plural agreement, it has to be remarked that, in the case of human controllers, this can be either masculine or feminine, depending on the biological sex of the referent, as examples (8) and (9) show:

8) Muḥammara (Khuzestan), male, 30 years (Leitner, own data)

əbbahāt-na ɡānaw yṣōfūn.
fathers-PRON.1PL be.PRF.3M.PL tell.PRS.3M.PL

“Our fathers used to tell.”

9) Ḥamidiyya (Khuzestan), female, 60 years (Leitner, own data)

əyan ən-nəswān.
come.PRF.3F.PL DEF-women

“The women came.”

M.PL. agreement, however, is more common, because whenever the controller denotes a group consisting of both males and females, or when said group is underspecified for gender, masculine agreement prevails over feminine. This is true for both plural controllers (examples 10 and 11) and collective ones (examples 12 and 13: note that here the controllers are the same that we saw attracting F.SG. agreement in examples 1 and 2):

8) Muḥammara (Khuzestan), male, 30 years (Leitner, own data)

əbbahāt-na ɡānaw yṣōfūn.
fathers-PRON.1PL be.PRF.3M.PL tell.PRS.3M.PL

“Our fathers used to tell.”

9) Ḥamidiyya (Khuzestan), female, 60 years (Leitner, own data)

əyan ən-nəswān.
come.PRF.3F.PL DEF-women

“The women came.”

M.PL. agreement, however, is more common, because whenever the controller denotes a group consisting of both males and females, or when said group is underspecified for gender, masculine agreement prevails over feminine. This is true for both plural controllers (examples 10 and 11) and collective ones (examples 12 and 13: note that here the controllers are the same that we saw attracting F.SG. agreement in examples 1 and 2):
10) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

əl-banāt u-l-ulād yībūn ḥdēm-āt-hum.

DEF-girls and-DEF-boys bring.PRS.3M.PL clothes.DIM-PL-PRON.3M.PL

“The girls and the boys bring their clothes.”

11) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

əl-ʔaḥwāziyyīn ygūlū(n) yōxān.

DEF-Ahwazi_people say.PRS.3M.PL-3M.SG.DAT storage

“The people of Ahwaz call it yōxān.”

12) Ahwaz (Khuzestan), male, 35 years (Leitner, own data)

yḥəbbūn-ḥa n-nās.

love.PRS.3M.PL-PRON.3F.SG DEF-people

“The people love it.”

13) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

ahl en-naxīl yḥuṭṭūn bī əl-tamur.

people DEF-palm_trees put.PRS.3M.PL in.PRON.3M.SG DEF-dates

“The date farmers put dates in it.”

If we now turn to nonhuman controllers, we see that they behave in a rather different way. M.PL. agreement, which was the most common type of agreement with human nouns, here disappears almost entirely, whereas F.PL. agreement becomes the dominant agreement type\textsuperscript{16}. This is consistent with all studies on agreement in gender-distinguishing varieties of Arabic, and with the theory put forward by Bettega (2019) and discussed in §1, according to which all nouns referring to nonhuman entities belong to agreement classes II or III and will, therefore, always attract feminine plural agreement when plural, rather than masculine\textsuperscript{17}. In the following examples we can see a representative item for each

\textsuperscript{16} This is clearly evident with non-collective controllers. Collective ones, due to their inherently low level of individuation, tend to attract F.SG. agreement, as we have seen. However, when they do attract plural agreement, it is always feminine plural, and not masculine.

\textsuperscript{17} Instances of M.PL. agreement occurring with nonhuman controllers are so rare that they can probably be dismissed as “slip of the tongue” on the part of the speakers, though some of them appear to be semantically motivated. One occurrence is found in the data of əbyūt, “houses” attracting M.PL. agreement, where the reference is clearly to the people and families who inhabit the actual buildings. Also əfīl, “male water buffalos” appears once triggering M.PL. agreement (see example (24) below): it is possible that biologically masculine domestic animals are to be considered as “borderline” items, oscillating between agreement classes I (human males) and III (nonhuman entities). Another example found in the data, ələb u-ṭōr marbūṭīn, “a dog and a bull tied-M.PL up”, would seem to confirm this hypothesis. This would not be
of the three agreement classes: in (14a) and (14b) the noun \( \text{walad} \) (pl. \( \text{awlād} \)) attracts M.SG. agreement when singular and M.PL. agreement when plural. This is a typical exponent of agreement class I, that only contains masculine human referents. In (15a) and (15b) \( \text{mara} \) (pl. \( \text{nəswān} \)) attracts F.SG. agreement when singular and F.PL. when plural: \( \text{mara} \) “woman” is an exponent of the (exclusively feminine) human members of agreement class II. This class also contains several nouns referring to nonhuman entities, as exemplified by \( \text{madārsa} \) (pl. \( \text{madāris} \)) in (16a) and (16b). In (17a) and (17b), finally, we see \( \text{yōm} \) (pl. \( \text{ayyām} \)) triggering M.SG. when singular but F.PL. when plural, as all members of agreement class III do (note that this class consists exclusively of nonhuman referents):

14a) Ahwaz (Khuzestan), female, 50 years (Leitner, own data)
\[
\text{hāḍa} \ l-\text{walad} \ yṛūḥ \ yəḥʕab. \\
\]
DEM.M.SG DEF-boy go.PRS.3M.SG play.PRS.3M.SG
“This boy goes to play.”

14b) Hamidiyya (Khuzestan), female, 60 years (Leitner, own data)
\[
\text{l-\text{ulād}} \ ə̇\text{kubraw} \ yə̇\text{wwəzēt-hum} \\
\]
DEF-children grow_up.3M.PL marry_off.PRF.1SG
“The children grew up and I married them off.”

15a) Ahwaz (Khuzestan), female, 50 years (Leitner, own data)
\[
\text{ḏīč} \ ə̇\text{mara} \ ʕad-\text{ha} \ fard \ ulēd. \\
\]
DEM.F.SG DEF-woman at-PRON.3F.SG INDEF child.DIM
“This woman, she only had one child.”

15b) Hamidiyya (Khuzestan), female, 60 years (Leitner, own data)
\[
\text{n-\text{nəswān}} \ ə̇\text{l} \ yə̇\text{diggan} \\
\]
DEF-women REL tattoo.PRS.3F.PL
“The women who were tattooing.”

16a) Ahwaz (Khuzestan), female, 40 years (Leitner, own data)
\[
\text{ḥāy} \ l-\text{madārsa} \ ə̇\text{ñe} \ ̣\text{garēt} \ bī-ha \\
\]
DEM.F.SG DEF-school REL 1SG study.PRF.1SG in-PRON.3F.SG
“This is the school in which I studied.”

typologically unprecedented, since in many languages controllers of this type enjoy “semi-human” status. However, data are too scarce to explore the phenomenon in detail.
16b) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

madārās-kum ntum hna lli tšūfūn-hən.

schools-PRON.2M.PL 2M.PL here REL see.PRS.2M.PL-PRON.3F.PL

“These schools of yours, which you see.”

17a) Ahwaz (Khuzestan), female, 50 years (Leitner, own data)

ḏak əl-yōm.

DEM.M.SG DEF-day

“That day.”

17b) Ahwaz (Khuzestan), female, 60 years (Leitner, own data)

rāhan ayyām w-əyyan ayyām.

go.PRF.3F.PL days and-come.PRF.3F.PL days

“Days went by and days came.”

Note that ayyām, “days” in example (17b) is the typical nonhuman referent characterized by very low individuation that one would expect to trigger F.SG. agreement. This actually happens in the very same expression (rāḥat ayyām u-yat ayyām, “Days came-F.SG. and went-F.SG. by”) in a text from a different speaker18. It has to be kept in mind (as has been remarked by several authors, see for instance Holes 2016, 326-354) that, although the individuation hypothesis is able to account for a great deal of the variation in agreement we observe in almost every variety of SA, exceptions can always be found. In actual language use, several pragmatic, stylistic and contextual factors can interfere with the final output. In general, as far as KhA is concerned, it is possible to note that percentages of F.PL. agreement with nonhuman referents appear to be particularly high (and percentages of F.SG. agreement particularly low) if compared to other gender-distinguishing varieties such as Bedouin Tunisian (Ritt-Benmimoun 2016), Omani (Bettega 2017) and Najdi Arabic (Bettega 2019). In 3.6 we will discuss how these percentages get even higher in the speech of the older generations of Khuzestanis.

3.2 Controller-related factors: concreteness

Another factor that has important consequences on agreement is the abstract or concrete nature of the controller. Since this distinction only applies to inanimate controllers, concreteness could be considered as a sub-factor of animacy.

As we have seen, non-individuated targets tend to have higher chances to attract (feminine) singular agreement than individuated ones. Individuation, as we discussed in §1, has to do with the ability of the human mind to focus on certain types of referents as the foci of attention. Human beings obviously rank very high

18 Note that Ritt-Benmimoun (2016: 277) as well reports the possibility of both F.SG. and F.PL. agreement with ayyām “days” in the Bedouin dialects of southern Tunisia.
in this respect, since they are endowed with the highest possible level of agentivity. Animals follow close by, then lifeless but mobile objects, and then inanimate and motionless entities such as landscape elements (some of which could be described as borderline elements in term of concreteness, since they have no clear limit or demarcation and/or are not tangible: think for instance of the sky, the clouds, a mountain range, and so on)\textsuperscript{19}. Purely abstract concepts stand at the bottom of the individuation hierarchy.

In light of the above, one would expect abstract controllers in KhA to attract higher percentages of singular agreement than their concrete counterparts. As table 4 shows, this is actually the case in our data, since abstract nouns appear to be almost twice as likely to trigger singular agreement than concrete controllers:

<table>
<thead>
<tr>
<th>Table 4 – Concreteness and agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singular targets depending on concrete controllers</strong></td>
</tr>
<tr>
<td>36.8% (70/190)</td>
</tr>
</tbody>
</table>

Although the percentage of F.SG. agreement with abstract nouns is very high in KhA, it is still possible to find a good number of these controllers attracting F.PL. agreement. This was the case with \textit{ayyām}, “days”, in example (17b) above, and it is also the case with \textit{məṣṭalaḥāt}, “expressions” and \textit{ḏikrayāt} “memories” in (18) and (19):

18) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

\textit{haγanni məṣṭalaḥāt maḥallīyyāt} tab\textit{Han maḥhad} yəfham-hən.  
DEM.F.PL expressions local.F.PL of course nobody understand.PRS.3M.SG-PRON.3F.PL

“This are local expressions, which of course no one understands.”

19) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

\textit{w-ạḥad ḏikrayāt-i aŋkar-hən…}  
and-one memories-PRON.1SG remember.PRS.1SG-PRON.3F.PL

“And one of the memories that I remember (is)…”

As already pointed out, it would seem that F.PL. agreement with members of class III is particularly common in KhA if compared to other Arabic dialects. If, as it

\textsuperscript{19} Hanitsch (2011: 146) refers to these “borderline” elements as “non-prototypical concrete nouns” (“nicht-prototypische Konkreta”), i.e. nouns referring to concrete entities and lacking one or more of the following traits: [+mobile] [+tridimensional] [+solid]). Unsurprisingly, she finds that in Damascus Arabic these nouns tend to attract F.SG. agreement.
seems to be the case, the occurrence of feminine plural agreement is to be considered an archaic feature in SA, then the Arabic of Khuzestan should be regarded as remarkably conservative in this respect.

3.3 Controller-related factors: quantification

All the existing studies on agreement in Arabic have found some type of correlation between agreement and quantification. In particular, numerals below ten have been found to almost systematically correlate with plural agreement. Our data confirm these findings and add some interesting detail to them. Since, as we have seen, the vast majority of human controllers in our data attract plural agreement, a statistical analysis of quantified vs. non-quantified human controllers would have been of little meaning (the percentages of singular agreement being negligible with both groups). Therefore, in this paragraph we will only consider the effects of quantification on nonhuman controllers (with no distinction between animate and inanimate ones).

Firstly, in table 5 we can see that a high degree of correlation exists between numerals below ten (including controllers inflected for dual number) and plural agreement:

<table>
<thead>
<tr>
<th>Singular Targets</th>
<th>Plural Targets</th>
<th>Total Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 (20%)</td>
<td>12 (80%)</td>
<td>15</td>
</tr>
</tbody>
</table>

Interestingly, all singular targets depend on dual controllers. While two of these targets are instances of the particle māl (that, as we have seen in fn. 14, is a partially grammaticalized element which not always inflects for gender and number), the remaining one is a verb in the feminine singular. Most studies on agreement and/or the dual number in SA (including Blanc’s 1979 seminal study on the topic) concord in maintaining that duals are treated as plurals when it comes to agreement, i.e. they systematically attract plural concord. Example (20) below shows that duals are indeed treated as plurals, to the point that they can even – under certain conditions – trigger feminine singular agreement, as real plurals can. In particular, the example features a verb preceding a controller with an intrinsically very low level of individuation (reinforced by the use of the adverb tağriban, “more or less”). Both lack of individuation and the pre-controller position of the target strongly favor F.SG. agreement, so that its appearance in this context is unsurprising:
Ahwaz (Khuzestan), male, 30 years (Leitner, own data)

More or less two months pass."

Controllers quantified by a numeral above ten are extremely rare in the data (only three targets depend on such controllers, two singular and a plural one). For this reason, we analyzed them together with other controllers quantified by elements that tend to stress the perception of the referent as an indistinct mass or group, or in general as a plurality whose exact number is not known. These quantifiers are kaṯīr and wāyəd, both meaning “many, a lot” (8 occurrences), ǧalīl “few, a little” (3 occurrences), ǧill “all” (2 occurrences), and aşgadd, “how many?” (1 occurrence), plus 3 different quantifiers all meaning “some” (although with different semantic nuances; these are ši, čam and furəd, 1 occurrence each). The results of our analysis are shown in table 6:

Table 6 – Agreement with controllers quantified by a vague quantifier (including numerals above 10)

<table>
<thead>
<tr>
<th>Singular Targets</th>
<th>Plural Targets</th>
<th>Total Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>23 (88.5%)</td>
<td>3 (11.5%)</td>
<td>26</td>
</tr>
</tbody>
</table>

Contrary with what happened with numerals below ten, vague quantifiers strongly favor singular agreement\(^{20}\), in that they subtract to the referent’s level of individuation.

3.4 Target-related factors: CT distance and target type

In his study on Cairene, Belnap (1991: 86) found distance between T(arget) and C(ontroller), expressed in terms of phonological words, to be the second most important variable affecting agreement after controller type. This factor has been often overlooked in subsequent studies on the topic. However, as we will see, it is of utmost importance to understand the functioning of agreement in SA in general, and KhA in particular.

It is important to bear in mind that distance between C and T is also mutually dependent on another factor, namely target type. Certain target types are naturally more prone to be found at a certain distance from their controller (e.g. right next to it, or very far away), and as a consequence, they are more likely to show certain types of agreement. Obviously, this reasoning could be reversed, stating that

\(^{20}\) Note also that in all varieties of Arabic a noun quantified by a numeral above ten is always singular (even though it obviously is a plural from a semantic point of view). This reinforces the idea that numerals above ten have collective-like properties. Also, the quantifier čam requires the noun following it to be in the singular.
certain target types are more likely to trigger a certain type of agreement, and because of this that specific kind of agreement is more often encountered at a certain distance from the controller. At present, it is not clear which one, among the two factors, is to be considered the dominant one (if a dominant one exists at all). In this paragraph, both phenomena will be treated together, covering CT distance first and then moving on to discuss target type.

Table 7 shows how the likelihood of singular agreement with plural/collective controllers tends to decrease the further away we move from the controller. In other words, increase in distance makes the target more likely to show plural agreement. Belnap (1991: 86) linked this fact to the notion of recoverability of information: the more linguistic material is inserted between two words, the more the semantic and grammatical relations between the two will become opaque (and hence the need to reinforce them via morphological means).

<table>
<thead>
<tr>
<th>CT Distance (in terms of phonological words)</th>
<th>Percentage of singular (M. or F.) agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48.6% (68/140)</td>
</tr>
<tr>
<td>2</td>
<td>31% (26/84)</td>
</tr>
<tr>
<td>3</td>
<td>25.5% (12/47)</td>
</tr>
<tr>
<td>4</td>
<td>25.7% (9/35)</td>
</tr>
<tr>
<td>5 or 6</td>
<td>18.1% (6/33)</td>
</tr>
<tr>
<td>7 or 8</td>
<td>20% (4/20)</td>
</tr>
<tr>
<td>9 or 10</td>
<td>11.1% (2/18)</td>
</tr>
<tr>
<td>11 and above</td>
<td>18.2% (8/44)</td>
</tr>
</tbody>
</table>

As can be seen, the decrease is not strictly monotonic, but the tendency of singular agreement to cluster in the proximity of the controller is nonetheless clear. It seems probable that the analysis of a much larger database would level out the few irregularities which our table presents (in particular the spike that can be observed in the targets occurring at a distance of more than ten words. Note, however, that singular targets disappear entirely after 17 words, and up to 33,

Note that targets appearing before their controllers have not been calculated in Table 7. The percentage of singular agreement among these targets is 43% (34 out of 79), and it raises to 50% in the case of TC verbs. Demonstratives, on the contrary, appear to favor plural agreement: only 28.6% of all demonstratives occurring in TC structures takes singular agreement (on this point see below). Note also that targets whose distance from the controller was impossible to determine have not been computed in Table 7. This was the case, for instance, of all those targets which appeared in the context of a dialogical interaction, which referred back to a controller previously mentioned by another speaker.

---

21 Note that targets appearing before their controllers have not been calculated in table 7. The percentage of singular agreement among these targets is 43% (34 out of 79), and it raises to 50% in the case of TC verbs. Demonstratives, on the contrary, appear to favor plural agreement: only 28.6% of all demonstratives occurring in TC structures takes singular agreement (on this point see below). Note also that targets whose distance from the controller was impossible to determine have not been computed in Table 7. This was the case, for instance, of all those targets which appeared in the context of a dialogical interaction, which referred back to a controller previously mentioned by another speaker.
which is the maximum distance observed in our corpus). CT distance appears to be responsible for most of the cases of agreement mismatch that appear in our data. Examples of this type are lengthy, so that, due to space constraints, we will only present two here. They both feature a collective controller, a human one in (21) and an inanimate one in (22). Note how in the second example the controller triggers three different types of agreement: first several M.SG. targets, then a F.SG. one and finally a F.PL one:

21) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

\[\text{gabul ma\text{f}alan an-n\text{\text{"a}}s alli str\text{\text{"a}}h barra b-\text{a}y-\text{a}y\text{r\text{"a}},}\]

before for_example DEF-people REL go.PRS.3F.SG out in-DEF-desert

“For example, in former times people used to go out to the desert,”

\[\text{\text{\text{"a}b}\text{\text{"a}n} l\text{\text{"a}} \text{\text{"a}d-hum} \text{tan}n\text{\text{"u}}r wa-l\text{\text{"a}} \text{\text{"a}d-hum} \text{\text{"u}.}}\]

of_course not at-PRON.3M.PL oven and-not at-PRON.3M.PL thing

“…they had of course no clay oven nor did they have anything.”

22) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

\[\text{fa-h\text{\text{"a}}d\text{\text{"a}} l-x\text{\text{"u}}, y\text{\text{"a}x\text{\text{"u}}n-a} w-ynazg\text{\text{"u}}n m\text{\text{"a}n} \text{\text{"o}nd-a,} \text{\text{"\text{"a}s-m-a}…}\]

so-DEM. DEF-palm_ take.PRS.3M.PL and-weave. from at-PRON. what-name-

M.SG leaves -PRON.3M.SG PRS.3M.PL 3M.SG PRON.3M.SG

“So these palm leaves, they take it and they weave with it, what is it called,

\[\text{ynazg\text{\text{"u}}n m\text{\text{"a}n} \text{\text{"o}nd-a} h\text{\text{"a}y} \text{al-...} \text{al-b\text{\text{"a}r\text{\text{"a}}}, yg\text{\text{"u}}l-la}\]

weave.PRS. from at-PRON. DEM.F.SG DEF DEF-mat say.PRS.3M.SG-

3M.PL 3M.SG 3M.SG.DAT

… they weave with it this, the… the b\text{\text{"a}r\text{\text{"a}}}, they call it,

\[\text{\text{\text{"a}s-m-a?} \text{al-ball\text{\text{"a}}al-ball\text{\text{"a}}, ysaww\text{\text{"u}}n m\text{\text{"a}n} \text{\text{"o}nd-ha} w-\text{\text{"a}b}\text{\text{"a}n} \text{\text{"u}.}}\]

what-name- DEF-mat DEF-mat make.PRS. from at-PRON. and-of_ even

PRON.3M.SG 3M.PL 3F.SG course

what’s its name? balla, balla, they make with it [i.e., with palm leaves],

and of course even

\[\text{fi \text{\text{"u}n\text{\text{"u}} \text{al-fi\text{\text{"a}r\text{\text{"a}}q} ham ysaww\text{\text{"u}}n [m\text{\text{"a}n}] \text{\text{"o}nd-hun} \text{al-ball\text{\text{"a}}.}}\]

in south DEF-Iraq also make.PRS.3M.PL from at-PRON. DEF-mat

3F.PL

… in the south of Iraq they make (it) with it the balla.”

Moving on to the relation between agreement and target type, this is the object of a well-known generalization in linguistics, called the “Agreement Hierarchy”. Corbett (2006: 207) states that: “For any controller that permits alternative agreements, as we move rightwards along the Agreement Hierarchy, the likelihood of agreement with greater semantic justification will increase monotonically”. The Hierarchy is illustrated in Table 8:
Table 8 – Corbett’s Agreement Hierarchy

|   |  
|---|---|
| attributive | predicate > relative pronoun > personal pronoun |

Given that KhA, as most varieties of SA, possesses an agreement option specifically dedicated to non-individuated agreement (F.SG.), it is of course debatable what “agreement with greater semantic justification” means in this context (since this largely depends on the controller we are considering; note that this problem was already highlighted by Belnap 1991: 88-9). In general, though, Corbett’s hierarchy appears to work well with our Khuzestani data, as shown in Table 9:

Table 9 – Target type and agreement

<table>
<thead>
<tr>
<th>Target type</th>
<th>Singular agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributive Adjectives</td>
<td>57.8% (26/45)</td>
</tr>
<tr>
<td>Predicative Adjectives</td>
<td>56.5% (13/23)</td>
</tr>
<tr>
<td>Preposition māl</td>
<td>50% (6/12)</td>
</tr>
<tr>
<td>Verbs</td>
<td>26.5% (41/155)</td>
</tr>
<tr>
<td>Pronouns</td>
<td>29.3% (53/181)</td>
</tr>
<tr>
<td>Demonstratives</td>
<td>0% (0/26)</td>
</tr>
</tbody>
</table>

As can be seen, decrease of singular agreement (and increase in plural) is almost monotonic as we move along the agreement hierarchy. The only elements that really stand out as odd are the demonstratives. Since demonstratives can be used either attributively or predicatively, one would expect them to pattern in between adjectives and verbs. On the contrary, they show an absolute preference for plural agreement (this preference is so strong that even in pre-controller position the vast majority of demonstratives attracts plural agreement, see fn. 21). This not only contradicts the agreement hierarchy, but is also in stark contrast with Belnap’s 1991 data on Cairene Arabic. One possible explanation for this remarkable discrepancy is that in KhA individuation has an even stronger impact on agreement than in Cairene: nouns which are qualified by a demonstrative tend to have an intrinsically very high level of individuation (since the speaker is referring to very specific entities), and this might cause these nouns to systematically attract plural agreement. This point obviously remains in need of further research.

---

22 As was the case for CT distance, targets appearing before the controller had been removed from this calculation. Note that relative pronouns in KhA do not inflect for gender or number, and so are not featured in Table 9.

23 KhA is typologically uncommon in that it features prepositions that can be inflected for number and gender (namely māl, “of, belonging to”). Since this element is mostly used attributively (as in marāčīb mālāt fīn, “ships (made) of-F.SG. clay”, “clay ships”), it makes sense that it patterns in a way similar to that of adjectives.
Example (23) demonstrates the use of SG for the adjective and PL for pronouns in two targets depending on the same controller (ḏəkrāyāt “memories”). Note that here target type seems to have a stronger effect on agreement than CT-distance (since one of the two plural targets occurs before the singular one):

23) Ahwaz (Khuzestan), male, 50 years (Leitner, own data)

\[
\begin{array}{llll}
\text{ta}b\text{tā}n & d\text{əkr}āyāt & hō\text{m}nā & kāfīrā & w-\text{bī-hōn} & mō\text{ʔslmāt}, \\
\text{of_c}ourse & \text{memories} & 3\text{-F.PL} & \text{many.F.SG} & \text{and-at-pron.3.F.PL} & \text{painful.F.PL} \\
\text{w-}\text{bī-hōn} & \text{faraḥ}, & \text{and-at-pron.3.F.PL} & \text{joy} \\
\end{array}
\]

“Of course, there are a lot of memories, and among them are painful ones as well as joyful ones.”

3.5 Target-related factors: word order

One last fundamental target-related factor remains to be discussed, namely precedence. Precedence is a well described phenomenon in typological literature (see Corbett, 2006: 180), and it has powerful effects on agreement, which manifest themselves consistently across languages. This is not to say that precedence is a relevant agreement condition in all languages: however, when it is, it consistently favors a specific type of outcome, namely singular agreement with plural controllers.

Generally speaking, the effects of precedence on agreement are a well-known and much discussed phenomenon in the literature on Arabic varieties. In particular, Brustad (2000: 67) wrote that “In VS sentences […] verb-subject agreement may be neutralized in most (if not all) forms of Arabic”25. Belnap (1991: 89) found a strong correlation between word order and agreement in his study on Cairene. Our KhA data would seem to corroborate these claims.

Generally speaking, of the 512 targets that constitute our corpus, 79 precede their controller: of these, 34 show singular agreement (be it feminine or masculine). In other words, 43% of all pre-controller targets are singular, while

---

24 This type of mismatch, where plural agreement precedes the singular, is highly uncommon, but it does occasionally occur also in other dialects. Cf. for Bahrain the example provided by Holes (2016: 353), where again a plural pronoun is followed by a singular adjective.

25 The literature on the effects of precedence in Arabic has focused mostly on verbal targets. We tried to run a separate analysis for all pre-controller verbs that appeared in our corpus, and it resulted that verbs have slightly higher chances of attracting singular agreement in TC structures than other target types, such as demonstratives. In particular, 50% of all TC verbs triggered singular agreement (19 out of 38), while the percentage was 36.6% in the case of other targets types (15 out of 41). The numbers, however, are small, so that a bigger corpus would be needed in order to obtain more reliable figures.
the percentage is lower for post-controller ones (31.4%, i.e. 139 targets out of 442). However, as we have seen, humanness, animacy and collectiveness have a stronger influence on agreement, so that these data are better analyzed taking into account these categories as well.

Firstly, we have separated collective controllers from non-collective ones. Target-Controller (TC) structures involving collectives are too rare in the data to allow further subdivisions based on humanness/animacy, so that all collectives have been analyzed together in table 10. Although the numbers are small, pre-controller targets appear to favor singular agreement much more than post-controller ones:

**Table 10 – Effects of precedence on collective controllers**

<table>
<thead>
<tr>
<th></th>
<th>Singular agreement (M. or F.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Targets preceding a collective controller</td>
<td>70.6% (12/17)</td>
</tr>
<tr>
<td>Targets following a collective controller</td>
<td>48.6% (69/142)</td>
</tr>
</tbody>
</table>

If we now move on to non-collective controllers, we see that this trend is confirmed. Table 11 shows the percentages of singular agreement in TC and CT structures involving human and nonhuman plurals (here as well, numbers were too small to further divide nonhumans into animate and inanimate ones, so that they have been lumped together):

**Table 11 – Effects of precedence on plural controllers**

<table>
<thead>
<tr>
<th></th>
<th>Singular agreement (M. or F.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Targets preceding a human plural controller</td>
<td>11.8% (2/17)</td>
</tr>
<tr>
<td>Targets following a human plural controller</td>
<td>4.2% (4/95)</td>
</tr>
<tr>
<td>Targets preceding a nonhuman plural controller</td>
<td>44.4% (20/45)</td>
</tr>
<tr>
<td>Targets following a nonhuman plural controller</td>
<td>32.8% (67/205)</td>
</tr>
</tbody>
</table>

Note that, of the few cases of M.SG. agreement with non-collective controllers, some actually appear within the context of a TC structure, as examples (24) and (25) show. Note how in both examples we have agreement mismatch between the TC target and the CT one:
3.6 Speaker-related factors: age groups

As we have seen in §2, linguistic habits in the Arabic-speaking community of Khuzestan can vary from one generation to another. In order to account for the age factor in the analysis of the agreement tendencies in our data, we divided the KhA corpus into two age-groups: one for speakers below the age of 45 and another for speakers above that.

Comparing these two data sets and their respective agreement behavior provided the following results: in most regards, we could not observe significant differences in the choice of agreement patterns between younger and elderly speakers. However, a distinction seems to exist in the realm of inanimate controllers: younger speakers chose F.SG. agreement with 36% of all targets depending on such controllers, while elderly speakers did so only in 20% of cases. Consequently, the younger generation shows comparatively less F.PL. agreement with inanimate heads: the younger age group chose F.PL. agreement in 45% of all targets depending on inanimate controllers as compared to 69% of the older age group.

This difference in agreement tendencies among the two age groups might be explained as an influence of MSA. The younger generation, which has more contact with MSA via the internet and television, could be more likely to adopt MSA structures (i.e. F.SG. agreement with all types of nonhuman controllers) due to a more frequent exposure to them.

Alternatively (or in parallel), the influence of other varieties of Spoken Arabic might be affecting agreement choices in KhA. Dialects such as Cairene Arabic (which has become one of the most well-known dialects throughout the whole
Agreement Patterns in Khuzestani Arabic

Arab world, thanks to its diffusion in the media), has been shown to employ F.SG. agreement more extensively than other dialects (Belnap 1991). For a discussion on the extra-linguistic factors that can affect agreement in Spoken Arabic, see Bettega (2018).

4. Concluding remarks
Due to its “peripheral” position, and subsequent relative isolation from the rest of the Arabic-speaking world, the dialect of Khuzestan appears to show a number of conservative traits, at least as far as agreement patterns are concerned. Percentages of feminine plural agreement with plural controllers, in particular, appear to be particularly high.

This is probably because of two reasons: The first one is that KhA remains more or less unaffected by the influence of neighboring varieties that have already lost gender distinction in the plural. Some of the other dialects that still preserve gender distinction in the plural are currently in the process of losing it due to contact with such varieties. Occurrences of M.PL. agreement where F.PL. would be expected are so low as to be statistically insignificant. This is in stark contrast with dialects such as the Bedouin varieties of Southern Tunisia described by Ritt-Benmimoun (2016), where the replacement of feminine forms with masculine ones seems to be fairly widespread.

The second reason is that percentages of F.SG. agreement with plural controllers appear to be particularly low in KhA. As has been noted by several authors, exposure to MSA through mass literacy and the modern media has made F.SG. agreement with nonhuman controllers increasingly more common in the spoken dialects (see Watson 2013: 213 for Sanaa, Bettega 2018 for Oman, Owens & Bani-Yasin 1987 for rural Jordan). In spite of the fact that F.SG. agreement with plural controllers has probably always been an option in all varieties of Arabic (since we find it virtually all over the Arabic-speaking world, and also in the oldest attestations of the written language), its frequency has undeniably increased in recent times. By contrast, in a dialect such as KhA, which is spoken in a country where MSA is not an official language, percentages of F.SG. agreement appear to be more modest. Regarding the use of F.SG. agreement as well, however, a generational gap seems to exist between elderly and younger speakers, the latter being more prone to the use of the F.SG. with nonhuman controllers due to the fact that – even in Khuzestan – MSA is now a part of everyday life, thanks to satellite television.

In paragraphs 3.1 to 3.5 we have seen how several formal and semantic factors can, to different extents, affect the final agreement output. These factors can be related to the target (i.e. distance between target and controller, position of the target relative to its controller, and target type) or to the controller itself (i.e. its being quantified, its being a collective, and its denoting a human, animate, inanimate or abstract referent). This last factor, in particular, is connected to the
hypothesis, put forward in Bettega (2019), that gender-distinguishing varieties of Arabic are to be regarded as displaying a system of three separate agreement classes (the third class consisting of all those nouns that attract M.SG. agreement in the singular but F.PL. agreement in the plural). This description seems to apply to our KhA data, where nonhuman plurals seem indeed to almost categorically require F.PL. agreement (no matter what type of agreement they trigger in the singular). M.PL. agreement, on the contrary, remains restricted to biologically masculine human controllers (that is, of the controllers belonging to agreement class I).

In general, Belnap’s (1993: 113) remark that agreement in Cairene Arabic “is not an all-or-nothing matter […] it is a matter of statistical tendency” would seem to apply to KhA as well. While the type of plural agreement that a given controller will attract would seem to be always predictable with almost total accuracy, oscillation between singular and plural agreement is not always easy to account for. This is because, as we have seen, this oscillation largely depends on the referent’s level of individuation. Individuation itself, however, is just a convenient umbrella-term referring to a variety of phenomena (some interrelated, some very different in nature). Most importantly, as Brustad (2000) has noted, individuation primarily relies on the speaker’s perception of the referent: such perception can vary in time, space, and, most obviously, from one speaker to another. Attempts at predicting what type of agreement a speaker would choose for a given noun in a given context are possible: the final output, however, depends on a huge number of competing and/or overlapping formal, semantic and pragmatic factors, so that a certain degree of variation will always persist. For this reason, agreement patterns in SA are best understood by looking at the big picture: as demonstrated in the preceding paragraphs, even though single isolated occurrences can sometimes show unexpected behavior, clear patterns will always emerge from the analysis of a statistically significant corpus of data.

Bibliography


