



# Parsing written language with non-standard grammar

## An eye-tracking study of case marking in Arabic

Andreas Hallberg<sup>1</sup> · Diederick C. Niehorster<sup>2</sup>

Published online: 8 June 2020  
© The Author(s) 2020

### Abstract

Morphologically marked case is in Arabic a feature exclusive to the variety of Standard Arabic, with no parallel in the spoken varieties, and it is orthographically marked only on some word classes in specific grammatical situations. In this study we test the hypothesis that readers of Arabic do not parse sentences for case and that orthographically marked case can therefore be removed with no effect on reading. Twenty-nine participants read sentences in which one of the two most frequent types of orthographically marked case was either retained or omitted, while their eye-movements were monitored. The removal of case marking from subjects in the sound masculine plural declension (changing the suffix *-ūn* *ون* to *-īn* *ين*) had no negative effect on gaze duration, regressions out, or go-past time. The removal of case marking from direct objects in the triptote declension (omitting the suffix *-an* *ان*) did however result in an increase in these measures. These results indicate that only some forms of case marking are required in the grammar used by readers for parsing written text.

**Keywords** Arabic · Case · Eye movements · Reading · Sentence processing

Standard written varieties of languages often have grammatical features not present in spoken varieties of the same language. Users of these languages therefore engage in two grammars, one for speech production in everyday situations and one for reading and writing (Carter, 1999; Trudgill, 1975). In cases where features specific to the standard variety are superfluous for sentence parsing it is possible that readers develop reading strategies where this feature is ignored and not registered as

---

✉ Andreas Hallberg  
andreas.hallberg@sprak.gu.se

<sup>1</sup> Department of Languages and Literatures, University of Gothenburg, Box 200, 40530 Göteborg, Sweden

<sup>2</sup> Lund University Humanities Lab and Department of Psychology, Lund University, Lund, Sweden

carrying grammatical information—over-ridden, as it were, by non-standard grammar used in parsing. In the present study we investigate one such situation: morphologically marked case in Arabic. As detailed below, case markers are considered a hallmark of correct Arabic, yet they are only occasionally present in writing and the syntactic information that they provide is redundant. In this study we investigate whether readers parse sentences using a “case-less” grammar, akin to that of the spoken varieties of Arabic, in which case marking is not expected, or whether they parse sentences with a grammar that requires case marking.

## Case in Arabic

Arabic is a prototypical case of *diglossia* (Ferguson, 1959, 1996), meaning that the linguistic varieties used in everyday spoken interaction differ significantly from the standard variety, Standard Arabic, which is used primarily in writing and is used orally only in formal, public situations (speeches, lectures, news broadcasts, etc.). The spoken varieties are acquired in childhood as first languages. Standard Arabic is acquired later through formal education as a second language, although children are exposed to it from a young age through audiovisual media (Saiegh-Haddad, 2003). The spoken varieties are also commonly used in informal written communication in text messaging and on social media (Abu Elhija, 2014; Kindt, Høigilt, & Kebede, 2016). The vast majority of writing is nevertheless in Standard Arabic, and readers have an advantage in visual word recognition in written Standard Arabic over written forms of their spoken variety (Nevat, Khateb, & Prior, 2014). One of the features specific to Standard Arabic is its system of morphologically marked case: a set of suffixes on nouns and adjectives marking their syntactic roles in the clause. None of the spoken varieties has a system of morphologically marked case parallel to that of Standard Arabic (Fischer & Jastrow, 1980; Versteegh, 2004), and to the extent that speakers acquire this system they do so as part of their acquisition of Standard Arabic.

There are three cases in Standard Arabic, nominative, accusative and genitive (Badawi, Carter, & Gully, 2004), but the morphological marking of these cases is largely absent in most written text. Case is most often marked phonologically with a word final short vowel, in indefinite nouns followed by /n/, called *nunation*. Short vowels and nunation are represented in the Arabic writing system with diacritics. Diacritics are, however, only used in a limited set of text types, most importantly children’s literature and religious source texts, while the default is for text to be undiacritized (Saiegh-Haddad & Henkin-Roitfarb, 2014), and thus lacking a representation for most forms of case marking. In the default undiacritized orthography, case marking affects the orthographic shape of the word, by adding or altering a letter, only in certain words in certain grammatical contexts. This type of case marking will here be referred to as *orthographically marked case*. Only orthographically marked case as it appears in the default undiacritized orthography is investigated in this study. Investigating other forms of case markers would necessarily entail experimental conditions where diacritics are added to texts, which are then not representative of normal Arabic text.

The two most frequent inflectional classes with orthographically marked case are the *sound masculine plural* (e.g., *mu'allimūn* ‘teachers’) and the *triptote* (e.g.,

**Table 1** Orthographically marked case in the sound masculine plural and triptote inflectional classes

	Sound masculine plural			Triptote		
	Standard Ar.		Spoken Ar.	Standard Ar.		Spoken Ar.
Nom.	<i>mu'allimān</i>	معلمون	<i>m'allimān</i>	<i>walad</i>	ولد	<i>walad</i>
Acc.	<i>mu'allimān</i>	معلمين	<i>m'allimān</i>	<i>waladan</i>	ولداً	<i>walad</i>
Gen.	<i>mu'allimān</i>	معلمين	<i>m'allimān</i>	<i>walad</i>	ولد	<i>walad</i>

*walad* ‘boy’). The case endings in these two inflectional classes as they appear in undiacritized text and in the spoken varieties are summarized in Table 1. As can be seen in the table, the written Standard Arabic form differs morphologically from the Spoken Arabic form only in the nominative for the sound masculine plural, and only in the accusative for the triptote. The sound masculine plural takes the ending *-ūn* (ون) in the nominative and *-īn* (ين) in the accusative and the genitive. The latter form is invariably used for these words in all spoken varieties and is commonly used in Standard Arabic unscripted speech, even when the nominative form is prescribed (Hallberg, 2016). The ending *-īn*, while commonly described as marking accusative/genitive case, will therefore be regarded as the default form for the purposes of this study, and as unmarked for case. Words in the triptote inflectional class take orthographically marked case only in the accusative, with the ending *-an* (ان), and only if indefinite and masculine. This orthographically marked accusative ending consists of the letter *alif* (ل), but it is often accompanied with the diacritic *an* (ا), even in otherwise undiacritized text. There are a few other situations, apart from these two, where case is orthographically marked, but they are much less frequent and will not be further discussed in this paper. In total, around 6% of all nouns and adjectives in a natural Standard Arabic text have orthographically marked case (Hallberg, 2016).

The scarcity of case marking in writing rarely results in problems for comprehension since the syntactic role of constituents are determined by word order and verb agreement (Beeston, 1970; Holes, 2004; Versteegh, 2004). Accordingly, the ability to use case markers is not correlated with reading comprehension (Abu-Rabia, 2001; Khaldieh, 2001; Parkinson, 1993). When Standard Arabic is extemporaneously spoken, as opposed to being read aloud, for example in panel discussions and news interviews, case endings are used only vary sparingly and inconsistently (Badawi, 1985; Hallberg, 2016; Meiseles, 1977; Parkinson, 1994). In spoken Standard Arabic, orthographic case markers, such as those described above, are used at higher rates than are case markers that lack orthographic representations. Case markers that have an orthographic representation are used at rates of around 50% in speech, with large variation between speakers, while the rates of usage of other case markers are much lower (Hallberg, 2016). Poor mastery of active use of the system of case marking even by highly educated speakers of Arabic has often been noted in the literature (Beeston, 1970; Ibrahim, 1983; Kaye, 1972; Saiegh-Haddad & Schiff, 2016). It has also been experimentally demonstrated by Parkinson (1993). We interpret these facts as indicating that case marking is an optional feature in oral production of Standard

Arabic. It has been suggested that speakers employ the language production system to predict and therefore more quickly process language input (Pickering & Garrod, 2006, 2013). Thus, according to this view, if readers use their system of language production, in which case marking is inconsistent or absent, to predict upcoming written input, forms that are unmarked for case will accord with the reader's prediction and not interfere with their parsing of the sentence.

## Prescriptive anomalies and parsing anomalies

For the purposes of the present study we need to make a distinction between two types of syntactic anomalies. On the one hand there are linguistic structures that are anomalous with regards to the prescriptive grammatical system of a language, and on the other hand there are linguistic structures that are anomalous with regards to the grammar used by the reader to parse the sentence. We propose the term *prescriptive anomaly* for the former and *parsing anomaly* for the latter. Prescriptive anomalies can be identified by comparing structures to authoritative grammatical descriptions of the standard language. Omitting the accusative marker *-an* from a direct object in Standard Arabic is an example of a prescriptive anomaly. Parsing anomalies, on the other hand, can only be detected by investigating readers' actual parsing of sentences containing these structures. If the omission of the accusative marker *-an* is not perceived by a reader as an anomaly and obstructs the reader's parsing of the sentence, it is only a prescriptive anomaly, not a parsing anomaly. We theorize that the omission of orthographically marked case, while being a prescriptive anomaly, does not constitute a parsing anomaly and will therefore not be noticed by readers.

An interesting point of comparison to orthographically marked case in Arabic is number agreement in English verbs, discussed in Pearlmutter, Garnsey, and Bock (1999). Like orthographically marked case in Arabic, gender agreement in English is syntactically redundant due to fixed word order; it is only marked in certain situations, namely in the third person singular in the present tense (*he/she/it reads* vs. *I/you/we/they read*), except for the verb *be*, which is always marked for number. Pearlmutter et al. (1999) speculate that "the comprehension system might be more efficient if it largely ignored agreement information, backtracking to handle it only when other constraints were insufficient." It is clear, however, that readers of English do not ignore number agreement, and that violations in number agreement therefore disrupt the reading process (Pearlmutter et al., 1999). Number agreement in English differs from orthographically marked case in Arabic, however, in being present in most spoken varieties of English. This most likely blocks readers from developing a parsing strategy that allows missing number agreement to go unnoticed. For orthographically marked case in Arabic, on the other hand, there is no such hindrance to a parsing strategy that ignores this feature, since it does not feature in the natively spoken Arabic varieties.

## Syntactic anomalies in the eye movement record

To test the hypothesis that the removal of orthographically marked case does not constitute a parsing anomaly we had participants read sentences from which case markers had been removed while their eye movements were recorded using eye-tracking. Eye tracking is a technique to study cognitive aspects of reading that has been used to investigate reading in a large number of languages (Rayner, 1998, 2009). In Arabic, eye-tracking has been used to investigate global eye-movement characteristics in text with and without diacritics (Chahine, 2012; Roman & Pavard, 1987), perceptual span (Jordan et al., 2013), processing of single words (Jordan, Almabruk, McGowan, & Paterson, 2011; Paterson, Almabruk, McGowan, White, & Jordan, 2015), and whether readers make use of grammatical information provided by diacritics (Hermena, Drieghe, Hellmuth, & Liversedge, 2015). This study is, to the best of our knowledge, the first eye-tracking study of the processing of case markers in Arabic.

Eye-tracking studies of reading syntactically anomalous sentences and garden-path sentences have consistently revealed that increased regressions from the anomalous word and from the next two or three words is indicative of difficulties in syntactic processing. Eye-tracking studies of syntactic processing in reading have primarily used garden-path sentences, where a possible initial interpretation of the sentence is contradicted by succeeding material in the sentence (e.g., Frazier & Rayner, 1982; Rayner & Sereno, 1994; for an overview, see Clifton, Staub, & Rayner, 2007). In (1), for example, *a mile and a half* is normally interpreted as the object of the first clause in first pass reading, an interpretation that must be altered in the later disambiguation region. In (1) the disambiguating region is the word *seems*. In the disambiguating region one typically finds longer fixation times and higher rates of regression compared to similarly constructed sentences where no garden-pathing occurs. Inflated reading times and increased rates of regression are thus indicators of disrupted syntactic parsing.

- (1) Since Jay always jogs a mile and a half seems like a very short distance to him. (Frazier & Rayner, 1982)

There has been relatively little research on eye movements in reading sentences where the processing is disrupted by an outright syntactic anomaly. The available studies do however report eye movement patterns similar to those found using garden-path sentences. Ni, Fodor, Crain, and Shankweiler (1998) investigated eye movements of participants reading sentences such as (2–4) where the verb is syntactically anomalous, as in (2), pragmatically anomalous, as in (3), or non-anomalous, as in (4).

- (2) It seems that the cats won't usually eating the food we put on the porch.  
 (3) It seems that the cats won't usually bake the food we put on the porch.  
 (4) It seems that the cats won't usually eat the food we put on the porch.

For syntactically anomalous sentences they found no effect on the first pass reading time anywhere in the sentence, but a sharp increase in regressions out from the region containing the anomalous word, an increase that disappears by the third word after the anomaly. For pragmatically anomalous sentences they found both increased first pass reading times and regressions out, beginning after the site of the anomaly but progressively increasing towards the end of the sentence. These results were reproduced in Braze, Shankweiler, Ni, and Palumbo (2002), with the difference that they also found an increase in first pass reading times at the site of the syntactic anomaly that then disappeared in the following region. They also performed a word-by-word analysis showing that the increase in regressions out, but not reading time, began directly at the syntactically anomalous word and continued to the third word after the anomaly. Several studies have used eye-tracking to investigate the attraction phenomenon: the erroneous agreement of a word with a distractor noun closer than the head noun (Dank, Deutsch, & Bock, 2015; Deutsch & Bentin, 2001; Pearlmutter et al., 1999). These studies include various combinations of syntactically anomalous stimulus sentences with a distractors, such as (5), where the verb *where* erroneously agrees with the closes noun *cabinets*, rather than the head noun *key*, and sentences without distractor, such as (6), that are more clearly anomalous.

- (5) The key to the cabinets were rusty from many years of disuse.  
 (6) The key to the cabinet were rusty from many years of disuse

These studies report increased regressions out and increased total reading times on the anomalous word in sentences without distractor (6), as compared to anomalous sentences with distractor (5) and to non-anomalous control sentences. Syntactic anomalies thus reliably produce increased regressions out from the site of the anomaly and from subsequent words, and often also longer reading time.

## This study

Summing up, orthographically marked case is a morphological feature that (a) provides syntactic information that is redundant for comprehension; (b) is only occasionally available; (c) is not represented in speakers' native variety; and (d) is not mastered by most skilled readers. Based on this description we hypothesize that skilled readers of Arabic parse orthographically marked case as an optional feature and that the removal of case markers therefore does not constitute a parsing anomaly. This would make the parsing system similar to the production system of spoken Standard Arabic, where case marking is optional, and would differentiate it from prescriptive grammar, where case marking is compulsory. We test this hypothesis by monitoring eye movements during reading of sentences from which case markers that are prescriptively required have been removed. If the hypothesis is true then we expect not to find any significant increase in reading

times and regressions for these sentences, compared to sentences where the prescriptively required case markers are included. If the hypothesis is false then we expect a significant increase in these measures.

## Method

### Participants

An original thirty-two participants were recruited in Gothenburg, Sweden. Two participants were excluded due to problems with calibration and a further participant was excluded due to scoring only 69.4% on comprehension questions that were a part of our design (see the “Procedure” section below), which was more than 1 SD below the mean. This left 29 participants in the final analysis. Of these 29 participants, 22 were females (mean age<sup>1</sup> 33.1, *SD* 10.3), and all were recent immigrants from Arabic speaking countries (27 Syrians, one Iraqi, and one Moroccan) who had arrived in Sweden on average 3.8 years (*SD* 1.7) prior to the experiment. All were native speakers of Arabic with completed secondary education in an Arabic speaking country. Twenty-three of the participants had also obtained tertiary education in an Arabic speaking country (avg. 3.6 years, *SD* 1.2). Informed consent was obtained from participants on arrival. All were naive to the purpose of the experiment and had normal or corrected-to-normal vision.

### Stimulus and apparatus

Stimulus sentences were gathered from news articles from the Al Jazeera and BBC Arabic news sites<sup>2</sup> and modified to fit the sentence structures described below. We chose to have stimuli representative of news discourse in order for participants to expect prescriptively correct text. Sentences were displayed in the “Simplified Arabic” font at 20-point font size with no line breaks. All stimulus sentences begin with a pre-region and end with a post-region. The pre-region consists of three words, either a three-word adverbial prepositional phrase or a two-word adverbial prepositional phrase followed by a temporal adverb. The post-region consists of between three and thirteen words. Each stimulus sentence was displayed in either an unaltered or an altered condition. In the altered condition the target word was manipulated to create a prescriptively anomalous sentence by having a grammatical marker removed. Stimulus sentences were of three types which will be referred to as GENDER-, SMP-, and TRI-type sentences, with 20 sentences of each type. The structure of these three sentence types is illustrated in (7)–(9), with target words in bold face. Note that while the transcribed and glossed examples are written left-to-right the Arabic text is written right-to-left. GENDER-type sentences (7) include manipulations of gender marking in subjects. This type of stimulus sentence was included

---

<sup>1</sup> One participant did not wish to reveal their age.

<sup>2</sup> [www.aljazeera.net](http://www.aljazeera.net), [www.bbcarabic.com/arabic](http://www.bbcarabic.com/arabic).

in order to provide eye-movement data for a syntactic anomaly where we do expect a strong effect on eye-movements, that is, where we assume that the prescriptive anomaly is also a parsing anomaly, since the non-standard spoken varieties of Arabic also require gender congruency between verb and subject. In GENDER-type sentences, the fourth word, directly following the pre-region, is an intransitive verb in the third person feminine singular past tense with the suffix *-at* (تت). The fifth word is the target word and the subject of the preceding verb. It is a singular, human, and definite noun. In its unaltered condition, the target word is marked as feminine with the suffix *-a* (ة), thus agreeing with the preceding verb. In the altered condition, this suffix is removed, producing the masculine form. It is a syntactic anomaly as the subject does not agree in gender with the verb. SMP-type sentences (8) were designed to test the effect on eye movements of the removal of the nominative case marker in the sound masculine plural. In these sentences the word following the pre-region, the fourth word, is an intransitive verb in the third person masculine singular past tense. The fourth is a definite sound masculine plural subject and the target word.<sup>3</sup> This word has the nominative marking suffix *-īn* (ون) in the unaltered condition and a prescriptively anomalous *-īn* (ين) in the altered condition. TRI-type sentences (9) were designed to test effects on eye movements of the removal of the accusative marker in indefinite triptotes. In TRI-type sentences the fourth word, the word directly following the pre-region, is a transitive verb in the past tense. The fifth word is a definite and animate noun and the subject of the preceding verb. The sixth word is the target word. It is an inanimate indefinite direct object in the triptote declension. This word has the case marking suffix *-an* (ن) in the unaltered condition, while in the altered condition this suffix was removed. Fixed or common phrases with the accusative marker, such as *la'iba dawran* 'play a role' or *sajjala hadafan* 'score a goal', were avoided in stimulus sentences.

#### (7) GENDER-type sentence

في كتاب جديد تكلمت الكاتبة/الكاتب عن أساليب السلطة السوفياتية  
*fi kitāb jadīd takallamat al-kātib-a/-Ø 'an 'asālib as-sulta as-sufiyūtiyya ...*  
 in book new talked.F DEF-author-F/-M about methods DEF-authority DEF-Soviet ...  
 'In a new book, the author talked about the methods used by the Soviet authorities ...'

#### (8) SMP-type sentence

حسب خطاب الرئيس رغب الليبيون/الليبيين في التدخل الأجنبي في البلاد  
*ḥasab xiṭāb ar-ra 'is raḡiba al-libiyy-ūn/-īn fi at-tadaxxul al-'ajnabiyy fi al-bilād*  
 according.to speech DEF-president wanted DEF-Libyan-PL.NOM/-PL in  
 DEF-intervention DEF-foreign in DEF-country  
 'According to the president's speech, the Libyans wanted a foreign intervention in the country.'

<sup>3</sup> In Standard Arabic, the verb is inflected for the singular in the default VS word-order, even if this subject is plural (Badawi et al., 2004).



## (9) TRI-type sentence

بعد تدهور الأسعار توقعت الحكومة عجزاً/عجز في ميزانية البلاد  
*ba 'd tadahwur al-'as 'ār tawaqqa 'at al-hukūma 'aǧz-an/-ø fi mizāniyyat al-bilād*  
 after fall DEF- prices expected DEF- government **deficit- AccJ- ø** in budget  
 DEF- country  
 'After the fall in prices the government expected a deficit in the national budget.'

The stimulus sentences were designed to be as simple as possible with regards to case assignment; all three sentence types follow the default VSO word order and the target word is directly preceded by the verb (GENDER- and SMP-type sentences) or the verb and the subject (TRI-type sentences). None of the target words are part of a complex nominal phrase with an adjective or a nominal possessor. The first three post-target words do not contain any anaphora or congruency referring to the manipulated grammatical category (gender or case) that may interfere with spillover effects. To ensure that none of the words in the critical word positions (4–5 in SMP- and GENDER-type sentences and 5–6 in TRI-type sentences) were infrequent and therefore difficult to process, words in these positions were checked for occurrence in Buckwalter and Parkinson (2011), a list of the 5000 most frequent words in written Arabic. All target words are at least three letters long to minimize the chance of skipping.

Monocular eye movements were recorded with a desk-mounted EyeLink 1000 eye-tracker at a sample rate of 1000 Hz. A head rest was used to reduce head movements. Stimuli were displayed on a Dell 1704FPV's computer screen at 1280 × 1024 resolution.

## Procedure

Participants were informed that they were to read sentences from news articles and were instructed to read each sentence for comprehension at their normal reading pace. Instructions were given both orally and in writing. Interaction with the participants, including the consent form, oral and written instructions, and debriefing, was done in Arabic. After instructions and camera setup a nine-point calibration and validation procedure was performed. Calibrations with an average error exceeding 0.5° were repeated until a calibration error below 0.5° was achieved. The mean calibration error, including subsequent re-calibrations within sessions, was 0.36° (*SD* 0.13) with a mean maximum error for individual calibration points of 0.74° (*SD* 0.24).

Before each trial, participants looked at a fixation point at the center of the screen. After 2 s a cross appeared at the right-hand side of the screen, and when a fixation was detected on the cross, the whole sentence appeared with the first letter positioned at the location of the cross. Participants were instructed to look at the bottom left corner of the screen after finishing reading a sentence. When a fixation was detected in this part of the screen, the sentence was removed from the screen. Each participant read a total of 115 sentences, of which the first five were practice sentences not included in the analysis. Sixty stimulus sentences, twenty of each sentence type, were displayed in random order, together with 50 filler sentences of comparable length and complexity to the stimulus sentences. Half of the twenty stimulus

sentences of each type were displayed in the altered condition, and the other half in the unaltered condition. For every other participant the conditions were inverted. Twenty-four of the stimulus sentences, eight of each type, and ten of the filler sentences, were followed by a yes-or-no comprehension question displayed on the screen. Questions and correct answers were balanced between sentence types and conditions. Participants answered ‘yes’ or ‘no’ by pressing the left or right arrow key on the keyboard, marking either the word ‘yes’ or ‘no’ on the display screen. They could change their answer after the initial key press. When satisfied with their answer, they pressed the space bar on the keyboard to continue to the next trial. Answers to comprehension questions were recorded. Halfway through the experiment participants had the opportunity to take a break, after which the eye-tracker was re-calibrated. Other calibrations were also performed during the experiment when needed, as determined by visually inspecting the gaze position on a secondary monitor during the experiment. The session lasted on average 26 min. After the session the purpose of the experiment was explained to participants and they were rewarded with a cinema ticket for their participation.

## Analysis

Fixation classification was performed by the EyeLink 1000 host software (version 4.594) with the “cognitive” configuration. Fixations on the target word and the first three post-target words were analyzed. For these four word positions, data is reported on (a) *gaze duration*, the single-word equivalent to first pass reading time, that is, the sum of all fixations on a word from the first fixation on that word until but not including the first subsequent fixation on another word; (b) *regressions out*, the proportion of saccades from the word during the first pass reading targeting a previous word in the sentence; and (c) *go-past time* (also known as *regression path duration*), the duration from first fixating on the word to fixating on a subsequent word in the sentence, thus including any rereading of previous parts of the sentence and re-fixations on the word prior to a fixation on a later word in the sentence. Gaze duration is often taken as a measure of lexical processing, and increased regressions out and go-past time as an indication of difficulties in higher level processing (Rayner, Pollatsek, Ashby, & Clifton Jr, 2012).

Fixations longer than two standard deviations above the mean (572 ms) and shorter than 80 ms were excluded, resulting in an exclusion of 8.4% of fixations. The effects of the alteration condition on the aforementioned eye-movement measures were tested separately for each sentence type and word position in a Linear Mixed-Effects Model using the *lme4*-package (Bates, Mächler, Bolker, & Walker, 2015) in the *R* statistical software (R Core Team, 2013). The *LmerTest*-package (Kuznetsova, Brockhoff, & Christensen, 2017) was used to extract *p*-values from regression models. We performed the analysis on log-transformed data for the numerical measures of gaze duration and go-past time, and on the logit of the binary data of regressions out. In all models, item and participant were included as random effects with a maximal effect structure (Barr, Levy, Scheepers, & Tily, 2013). Where models did not converge, we removed random effects until convergence was achieved, first by removing correlation

between participant intercept and slope, then participant slope. This was only done in the model for regressions out, as noted in the results section. For go-past time, we also computed Bayesian factors for the null-hypothesis.

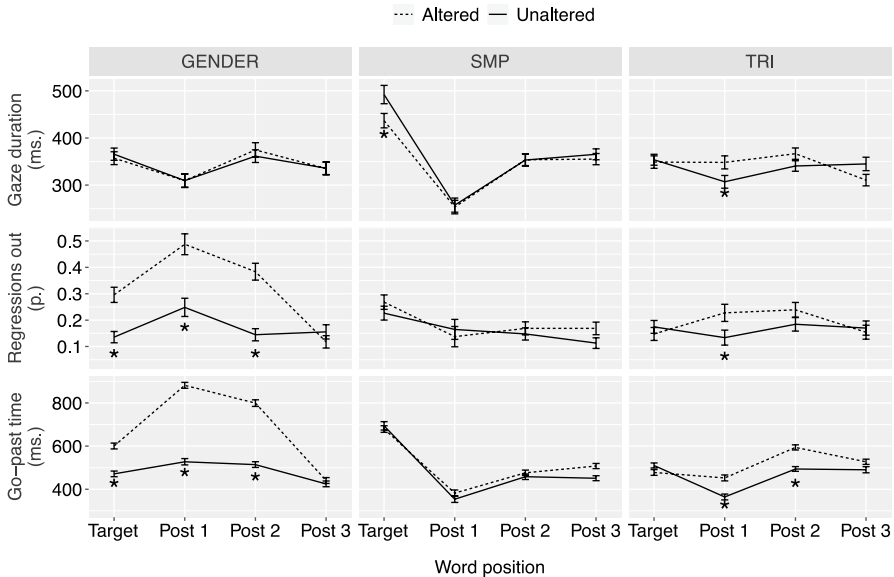
In Bayesian statistics, two models are compared with regards to their complementary probability in producing the observed data (Morey, Romeijn, & Rouder, 2016). The strength of the evidence in Bayesian statistics is measured in the form of Bayes factor, which is the ratio of the probability of one hypothesis to predict the data, relative to some alternative hypothesis. A Bayes factor of 2 for hypothesis A means that the data is twice as likely to occur under this hypothesis than under an alternative hypothesis B. The Bayes factor thus allows us to assess evidence in favor of a null-hypothesis, which is not possible in frequentist statistics (Kass & Raftery, 1995; Rouder, Speckman, Sun, Morey, & Iverson, 2009). We computed the Bayes factor for the null-hypothesis ( $BF_{01}$ ) using a Bayesian  $t$ -test (Rouder et al., 2009) performed with the *BayesFactor*-package in *R* (Morey & Rouder, 2018) in the three sentence types using the same maximal random effects structure as in the frequentist analysis presented above. We first computed  $BF_{01}$  for each sentence type in the target word and in the first and second post-target words. These are the word positions where any affect has been found in at least one of the three sentence types in the other analyses. We then computed  $BF_{01}$  for each sentence type with these three word positions together and with word position as an additional random effect.

## Results

Participants gave on average 88.5% correct answers to the comprehension questions ( $SD$  5.3), including questions following fillers. One participant scored more than  $1SD$  below the mean, at 69.4%. Data from this participant was removed from further analysis (as mentioned under “Participants” above). For stimulus trials, and with the low scoring participant excluded, participants gave on average 90.4% and 91.0% correct answers for sentences in the unaltered and altered conditions respectively. The difference is not significant ( $p = .81$ ). 8.4% of target words, 50.4% of first post-target words, 13.5% of second post-target words, and 22.7% of third post-target words were skipped on first pass. Descriptive data for gaze duration, regressions out, and go-past time is plotted in Fig. 1. Results from the regression models are listed in Table 2.

### Gaze duration

Average gaze durations are shown in the top row in Fig. 1. There was no difference between the two conditions in any of these word positions in GENDER-type sentences. In SMP-type sentences, interestingly, the only significant difference was a 55 ms *shorter* gaze duration on the target word in the prescriptively anomalous altered condition ( $t = -2.02, p < .001$ ). In TRI-type sentences the altered condition yielded a 41 ms longer gaze duration on the first post-target word ( $t = 2.54, p = .011$ ), with no other significant differences. The effect of the alteration condition was thus delayed one word in TRI-type sentences.



**Fig. 1** Descriptive data for gaze duration, regressions out, and go-past time in GENDER, SMP, and TRI-type sentences. Error bars indicate standard errors and asterisks significant differences ( $p < .05$ )

## Regressions out

Average proportions of regressions out are shown in the middle row in Fig. 1. The maximal regression models for regressions out failed to converge. Only random intercepts of participant and item were included in the models for TRI and SMP-type sentences, and only the random intercept of participant for GENDER-type sentences was included. In GENDER-type sentences there was a large and immediate difference in rate between the conditions in the target word and in the first two post-target words of .16, .24, and .24 respectively ( $z$ s  $> 4.4$ ,  $p$ s  $< .001$ ), and this difference had disappeared by the third post-target word ( $z = -1.06$ ,  $p = .288$ ). In SMP-type sentences there was no significant difference in regressions out in any of the word positions. In TRI-type sentences there was a significant difference in rate of .10 in the first post-target word ( $z = 2.0$ ,  $p = .046$ ). This pattern is similar to that found for gaze duration in TRI-type sentences in that there is only a difference on the post-target word, but not on the target word itself.

## Go-past time

Average go-past times are shown in the bottom row in Fig. 1. In GENDER-type sentences there was, similarly to regressions out, a large difference between the conditions for the target word and the first two post-target words, 129 ms, 355 ms, and 128 ms respectively ( $t$ s  $> 2.5$ ,  $p$ s  $< .01$ ). In SMP-type sentences there was no difference in any of the word positions. In TRI-type sentences there was no difference

**Table 2** Differences between alteration conditions

Target	Post 1			Post 2			Post 3									
	<i>b</i>	<i>SE</i>	<i>t/z</i>	<i>p</i>	<i>b</i>	<i>SE</i>	<i>t/z</i>	<i>p</i>	<i>b</i>	<i>SE</i>	<i>t/z</i>	<i>p</i>				
<i>Gaze duration (log)</i>																
GENDER	-.02	.04	-.47	.639	.01	.06	.24	.809	.00	.05	.09	.925	-.03	.05	-.48	.629
SMP	-.09	.04	-2.02	.044	.03	.07	.39	.695	.01	.04	.15	.884	-.03	.04	-.75	.450
TRI	-.04	.04	-.93	.351	.14	.05	2.56	.011	.04	.04	.92	.356	-.08	.05	-1.59	.113
<i>Rate of regression (logit)</i>																
GENDER	1.03	.23	4.42	<.001	1.12	.25	4.48	<.001	1.38	.23	5.88	<.001	-.33	.31	-1.06	.288
SMP	.22	.21	1.06	.288	-.17	.43	-.40	.689	.14	.26	.55	.585	.48	.26	1.83	.068
TRI	-.23	.26	-.88	.378	.63	.32	2.00	.046	.37	.24	1.54	.124	-.10	.28	-.34	.733
<i>Go-past time (log)</i>																
GENDER	.16	.06	2.65	.009	.48	.08	6.33	<.001	.38	.06	5.95	<.001	-.04	.07	-.53	.597
SMP	-.04	.05	-.76	.447	.05	.09	.58	.564	.04	.05	.79	.431	.06	.05	1.22	.222
TRI	-.08	.06	-1.35	.179	.20	.07	2.74	.006	.13	.06	2.08	.039	-.01	.07	-1.12	.902

**Table 3** Bayes factors for the null-hypotheses of the alteration condition on go-past time

Type	Target	Post 1	Post 2	All
GENDER	.1	<.001	<.001	<.001
SMP	7.1	5.0	6.6	14.7
TRI	2.5	.01	1.1	2.6

in the target word but only in the first and second post-target words, of 130 ms and 141 ms respectively ( $t = 2.75, p = .006$  and  $t = 2.08, p = .036$ ). We hypothesized that the removal of case markers in SMP- and TRI-type sentences would not affect the reading measures associated with difficulties in syntactic processing. The frequentist statistical models of go-past time presented above show that, at least for SMP-type sentences, we cannot reject the null-hypotheses that is in line with our prediction.

In order to assess evidence favoring the null-hypothesis, rather than only testing for its rejection, we conducted Bayesian statistical analyses on the log transformed go-past time. The computed Bayes factors are listed in Table 3.

In GENDER-type sentences, where the feminine marker was removed in the altered condition, the  $BF_{01}$  is 0.111 on the target word, and  $< .001$  on the first and second post-target words, as well as for all word positions taken together. There is thus decisive evidence against the null-hypothesis in GENDER-type sentences,<sup>4</sup> which is in accordance with our prediction. For SMP-type sentences,  $BF_{01}$  ranges from 5 to 7 in the three word positions, and is 14.7 for all three positions taken together. This provides strong evidence in favor of the null-hypothesis in SMP-type sentences, again in accordance with our prediction. For TRI-type sentences, Bayes factors are inconclusive for target words and second post-target words. For the first post-target word, the  $BF_{01}$  is .01, making the hypothesis of an effect 100 times more likely than the null-hypothesis for this specific word position. The  $BF_{01}$  of all three word positions together is inconclusive, giving negligible support for the null-hypothesis in TRI-type sentences. This provides localized strong evidence against the null-hypothesis in TRI-type sentences, which is contrary to our prediction.

## Discussion

In this study, we investigated whether skilled readers of Arabic parse orthographically marked case in Standard Arabic text as an optional grammatical feature. Our hypothesis was that they do, and that the omission of orthographically marked case from sentences would therefore not constitute a parsing anomaly and would not produce altered reading behavior as revealed by eye movements. We found support for this hypothesis in one of the two forms of case marking investigated.

In the experiment, participants also read sentences where the feminine marker had been removed from subjects of feminine inflected verbs. These sentences were included in order to produce eye-tracking data on a manipulation that is

<sup>4</sup> For evaluation of Bayes factors, see Kass and Raftery (1995).

assumed to constitute both a prescriptive anomaly and a parsing anomaly. As expected, the removal of the feminine marker in the altered condition produced patterns of eye-movements similar to those of previous eye-tracking studies on syntactic anomalies, namely no or only a small effect on the early measure of gaze duration and strongly increased regressions out and go-past time, lasting for two to three words from the anomaly (Braze et al., 2002; Dank et al., 2015; Deutsch & Bentin, 2001; Pearlmutter et al., 1999). These results from GENDER-type sentences demonstrate that our paradigm is able to identify parsing anomalies for the current language and population.

The other two sentence types in the experiment, SMP- and TRI-type sentences, were designed to directly address the hypothesis. In SMP-type sentences we changed the ending *-ūn* (ون) of sound masculine plural subjects to *-īn* (ين), thus removing nominative marking and creating a sentence of a form that is prescriptively anomalous, but corresponds to the morphology of vernacular Arabic and that is also common in extemporaneously spoken Standard Arabic. There was no indication that the alteration had any negative effects on reading in any of the eye-movement measures. A Bayesian statistical analysis furthermore yielded strong evidence in favor of the hypothesis that reading behavior was unaltered. It is worth reiterating here that case assignment in the stimulus sentences is as basic as it gets; the manipulated word is the subject of a directly preceding intransitive verb. In other words, we made it as easy as possible for participants to check for case marking. The lack of an effect of the removal of the case marker is thus not due to the stimulus sentences being grammatically complex. The fact that the omission of orthographically marked case in this construction had no negative effects on reading makes it highly unlikely that the same alteration would have negative effects on reading in more involved grammatical structures with, for example, long-distance case governance or less common forms of case assignment. We interpret the results presented here as strong evidence that readers do not check for orthographically marked case in the sound masculine plural, showing that case marking in these words is treated as optional. Furthermore, we cannot exclude the possibility that readers not only parse case marking in the sound masculine plural as optional, but that they also parse them as entirely unrelated to syntax. Such a theory would predict that not only the omission of case marking, but also the addition of prescriptively incorrect case marked forms, does not constitute a parsing anomaly.

The data for TRI-type sentences present a different picture. These sentences were altered by omitting the accusative marker *-an* (ن) from a triptote indefinite direct object. As with the manipulation of SMP-type sentences, this creates a sentence of a form that prescriptively anomalous, but corresponds to vernacular Arabic and is also common in extemporaneously spoken Standard Arabic. However, as opposed to the manipulation in SMP-type sentence, this did affect the reading measures, contrary to our expectations. Omitting the accusative marker resulted in increased gaze duration, regressions out and go-past time. These effects, however, were only found on the first post-target word, not on the altered target word itself. This effect is relatively small and highly localized, as compared with the effects of gender disagreement presented here, as well as when compared to other reports of the effect of syntactic anomalies found in the literature. Nevertheless, this indicates that the case

marking on indefinite triptote accusatives is not parsed as optional, but as a compulsory feature. The two forms of case markers investigated here are thus parsed differently. We are unsure as to why there is such a difference in the parsing of these two types of case marking, given the fact that both of these forms of case marking are absent in vernacular Arabic, so their removal would be expected to have a similar effect. One possibility is that the accusative ending on triptotes, being word final and often accompanied with a diacritic, is more visually salient than the case marked forms in the sound masculine plural, and is therefore more efficiently acquired on the basis of the visual input of printed text.

The effect of the omission of the case marking in TRI-type sentences was delayed one word on all three measures. A likely explanation for the one-word delay is the fact that in these sentences, the unmarked form of the target word used in the altered condition is prescriptively licensed if it is the first term in a genitive construction, that is, if it is followed by a nominal possessor. In Standard Arabic, if a noun is the first term in a genitive construction it does not take nunation. For a triptote noun in the accusative this means that rather than taking the suffix *-an*, which is orthographically represented, it takes the suffix *-a*, which is not, making it unmarked for case in undiacritized text. This is illustrated in (10) where the triptote direct object *qarār* 'decision' is marked for accusative with the ending *-an*, as is prescriptively correct. If this ending is removed, as in (11), the sentence is prescriptively incorrect. This is what was done in the altered condition in TRI-type sentences in this experiment. However, if the direct object is followed by a possessor noun in a genitive construction it does not take nunation and the case marker is therefore not orthographically represented (12). In the stimulus sentences, the target word was never followed by a noun, so that sentences such as (12) never occurred. This does mean, however, that if the reader does check for case marking, the anomaly caused by the removal of the accusative marker is detected only once the subsequent word has been identified, thus delaying any effect of the anomaly by one word.

- (10) دعم الوزير قراراً في البرلمان  
*da 'ama l-wazīr qarār-an fī l-barlamān*  
 supported DEF-minister decision-ACC in DEF-parliament  
 'The minister supported a decision in parliament.'
- (11) دعم الوزير قرار في البرلمان  
 \* *da 'ama l-wazīr qarār fī l-barlamān*  
 supported DEF-minister decision in DEF-parliament  
 'The minister supported a decision in parliament.'
- (12) دعم الوزير قرار الرئيس في البرلمان  
*da 'ama l-wazīr qarār ar-ra'īs fī l-barlamān*  
 supported DEF-minister decision DEF-president in DEF-parliament  
 'The minister supported the decision of the president in parliament.'



The removal of case marking in sound masculine plural subjects in SMP-type sentences (e.g., *siyāsīyy-īn* سياسيين ‘politician- PL’ instead of *siyāsīyy-ūn* سياسيون ‘politician- PL.NOM’) resulted in reduced gaze duration, which indicates facilitated lexical access (Rayner et al., 2012). This happens despite these words constituting a prescriptive syntactic anomaly in the stimulus sentences. The longer gaze duration observed on sound masculine plural nouns in the nominative form than in the form that is unmarked for case may be due to the fact the nominative marking morpheme *-ūn* does not exist in any of the spoken Arabic varieties and is thus absent in the first language of the reader. The nominative marked form can thus only be accessed through the Standard Arabic second language morphological system, which is presumably slower, due to familiarity and age of acquisition effects (Juhasz & Rayner, 2003, 2006). We did not, however, observe a similar facilitating effect of the removal of the case marker in TRI-type sentences, even though the case marked form (e.g., *walad-an* ولد ‘boy- ACC’) is also accessed through the Standard Arabic second language morphological system. This may be because any effect of facilitated lexical access of the unmarked form was hidden by the effect of the syntactic anomaly that was observed in these sentences.

## Conclusion

Our hypothesis that proficient readers of Arabic parse orthographically marked case as an optional feature was partially confirmed. The removal of orthographically marked case from sound masculine plurals had no effect on the eye movement record that could indicate parsing difficulties. The removal of the accusative ending from triptote direct objects did negatively effect reading. These two forms of case marking are thus parsed differently: one as optional and the other as compulsory. The omission of nominative case marking from sound masculine plural is a clear case of difference between prescriptive grammar and the grammar used by readers to parse the sentence. Readers parse these words with a nonstandard case-less grammar that over-rides standard, prescriptive rules of case marking.

**Acknowledgements** Open access funding provided by University of Gothenburg. The authors gratefully acknowledge Lund University Humanities Lab where extensive piloting was conducted.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix: Stimulus sentences

A slash separates the unaltered and the altered form of the target word.

### GENDER-type sentences

1. في مؤتمر للمحافظين تقدمت الوزيرة/الوزير للترشح للرئاسة وبالتالي لمنصب رئاسة الحكومة
2. في مقابلة أمس تحدثت الناطقة/الناطق باسم منظمة العفو الدولية عن الولايات المتحدة وانتقدت سياستها بشدة
3. في باريس اليوم أعربت متحدثة/متحدث باسم وزارة الخارجية الفرنسية عن قلق فرنسا من القانون الإسرائيلي الجديد
4. في نفس الوقت تساءلت الدكتورة/الدكتور في الاقتصاد عن مصداقية المعلومات الجديدة عن أرباح شركة أبل
5. قبل سبع سنوات تعرضت الممثلة/الممثل للتحرش الجنسي من قبل المنتج السينمائي المشهور
6. في العمل الجديد شعرت السيدة/السيد بأن المدير كان يبذل كل ما في وسعه لإجبارها على ترك وظيفتها
7. خلال رحلة سياحية وقعت الراكبة/الراكب من المركب في الماء الباردة وبقيت في الماء ساعات حتى جرى إنقاذها
8. من خلال المشروع تمكنت الباحثة/الباحث من الإثبات بأن الكربن يمكن أن يوقف خطر سرطان الأمعاء
9. في نيوزيلندا الأربعم ذهب الوزيرة/الوزير لشؤون المرأة بدراسة إلى المستشفى لتضع مولودها الأول
10. على هذا السؤال أجابت البروفيسورة/البروفيسور إنه يمكن مواجهة ألم العضلات من خلال تناول المزيد من البروتين
11. في كتاب جديد تكلمت الكاتبة/الكاتب عن أساليب السلطة السوفياتية في تضليل الناس.
12. في ألمانيا اليوم أشارت الرئيصة/الرئيس في خطاب إلى الرؤساء التنفيذيين إلى ضرورة توظيف المزيد من اللاجئين
13. بعد حفلة الافتتاح عبرت المديرية/المدير للدار الثقافي عن فرح المنظمة بالحضور الكبير في الاحتفال
14. حسب استطلاع الرأي تقدمت المرشحة/المرشح للرئاسة من الحزب الديمقراطي على منافسها الجمهوري
15. في المعتقل الإسرائيلي تعرضت الصحفية/الصحفي للتهديدات بأن الإنكار لن يفيدها وأنها ستدفع ثمن صمودها.
16. للمرة الرابعة بالتوالي عادت الفنانة/الفنان إلى مسرح مهرجان قرطاج الدولي
17. في تمهيد التقرير أشارت المديرية/المدير لقسم الشرق الأوسط إلى أن السلطات ألقت القبض على أكثر من عشرين فناناً
18. في الثلاثاء الماضي اختفت البريطانية/البريطاني بظروف غامضة خلال إقامتها في مدينة الغردقة
19. بعد الانتخابات التمهيدية حصلت الفلسطينية/ال فلسطيني على مقعد في الكونغرس حيث تمثل الحزب الديمقراطي
20. عبر التواصل الاجتماعي تحدثت اللاعبه/اللاعب عن الاختبارات للكشف عن المنشطات

## SMP-type sentences

21. في تركيا اليوم بدأ الممثلون/الممثلين عن الجمعيات الإغاثية بعملهم في المخيمات
22. في مؤتمر صحفي أجاب المسؤولون/المسؤولين عن أسئلة بشأن موت الوزير
23. في تصريح جديد رد السعوديون/السعوديين على النقد الموجه إليهم من منظمات حقوق الإنسان
24. منذ فترة طويلة يتكلم المهتمون/المهتمين في مجال الإعلان عن ضرورة التغيير في موقع تويتر
25. حسب خطاب الرئيس رغب الليبيين/اللبيين في التدخل الأجنبي في البلاد
26. في الهجوم الأخير اعتمد المعارضين/المعارضون على مختلف الأسلحة الثقيلة
27. خلال السنتين الماضيتين تمتع الإسلاميون/الإسلاميين في البلد بدعم كبير على مواقع التواصل الاجتماعي
28. في مساء الاثنين وصل المتظاهرون/المتظاهرين إلى باب السفارة الإسرائيلية
29. خلال الشهر الماضي خرج المعلمون/المعلمين في مظاهرات يومية يطالبون برفع رواتبهم
30. بعد قمة الاتحاد عبّر المشاركون/المشاركين فيها عن خيبة أملهم في الوصول إلى اتفاق
31. بعد التطورات الأخيرة انسحب المقاتلون/المقاتلين إلى قلب مدينة حلب في سوريا
32. بعد إبرام الاتفاق أتى المستثمرون/المستثمرين من دول عدة إلى إيران ليراقبوا سوق النفط
33. حسب التقرير الجديد قام المتشددون/المتشددين بشن هجوم كيميائي بسيط بالصواريخ.
34. في الأسبوع الماضي قام السياسيون/السياسيين بوقف بث برنامج تلفزيوني
35. إثر استقالة المدير اختلف المساهمون/المساهمين في جدوى الاتجاه الجديد الذي اتخذته الشركة
36. قبل ثلاثة أسابيع عاد الكويتيون/الكويتيين إلى المستوى المعتاد في إنتاج النفط
37. في جنوب أفغانستان اقترب المسلحون/المسلحين من مقر الصليب الأحمر في المنطقة
38. بعد الأحداث الأخيرة طالب المتخصصون/المتخصصين في الدستور بإلغاء إحالة المدنيين إلى محاكمات عسكرية
39. قبل أربعة أيام اختفى الموظفون/الموظفين في منطقة كانت قد شهدت العديد من عمليات الاختطاف
40. في الكنيست الإسرائيلي استمع المنتخبون/المنتخبين إلى خطاب رئيس الوزراء الجديد

## TRI-type sentences

41. بعد تدهور الأسعار توقعت الحكومة عجزاً/عجز في ميزانية البلاد
42. في القمة الدولية عقد الاتحاد اتفاقاً/اتفاق مع تركيا لتزويدها بثلاثة مليارات يورو
43. في التلفزيون المصري أثار المذيع غضباً/غضب حين تفوه بعبارات مسيئة للنساء
44. في دراسة علمية فحص العلماء أسماكاً/أسماك في البحر الأطلسي للكشف عن كمية البلاستيك في أجسامها
45. حسب موقع الجزيرة سمع الشهود انفجاراً/انفجار وعندما قدموا إلى الموقع رأوا قتيلين اثنين
46. بعد التصريحات الجديدة ينتظر الرئيس جواباً/جواب من بشار الأسد على التحذيرات التي وجهها لسوريا
47. حسب وزير الخارجية طالبت الوزارة تفسيراً/تفسير من السعودية بشأن خطط مشاركتها في العملية العسكرية
48. في تصريح رسمي أعلن الوزير تخفيضاً/تخفيض لإنتاج النفط من أجل تحديد التقلبات في الأسعار
49. في المحكمة العسكرية أصدر القاضي حكماً/حكم بالسجن المؤبد على أكثر من مئة متهم
50. حسب مصدر سعودي اتخذ الملك عدداً/عدد من القرارات وقام بمراجعة شاملة للعلاقات الخارجية
51. بعد اكتشاف الجثة فتح الادعاء تحقيقاً/تحقيق في جريمة قتل الطالب الإيطالي
52. خلال العقود الأخيرة شهد الشعب حروباً/حروب بالوكالة مما أعاق التطوير الاقتصادي
53. حسب مصدر سوري واجهت المعارضة هجوماً/هجوم من قوات النظام وأجبرتها على الانسحاب
54. في هجوم جوي قصف الروس هدفاً/هدف في ريف حلب مما أدى إلى قتل عشرة أشخاص
55. في النمسا اليوم اقترحت الحكومة تحديداً/تحديد لعدد طلبات الهجرة التي تقبلها
56. في باريس أمس قتلت الشرطة رجلاً/رجل باطلاق نار بعد أن قام بطعن شخصين بسكين
57. في عاصمة قطر وقع الأمير اتفاقاً/اتفاق مع الرئيس رجب طيب أردوغان لدعم الاقتصاد التركي
58. للحج هذه السنة دفع الفرد رسوماً/رسوم يتراوح مبلغها ما بين ثلاثة وأربع ألف ريال سعودي
59. في الدوري الإسباني رفض اللاعبون قراراً/قرار بأن تقام إحدى المباريات في الولايات المتحدة
60. على موقعهم الإلكتروني نشرت المنظمة تسجيلاً/تسجيل تقول إنه لصوت زعيمهم أبو بكر البغدادي.

## References

- Abu Elhija, D. (2014). A new writing system? Developing orthographies for writing Arabic dialects in electronic media. *Writing Systems Research*, 6, 190–214. <https://doi.org/10.1080/17586801.2013.868334>.
- Abu-Rabia, S. (2001). The role of vowels in reading Semitic scripts: Data from Arabic and Hebrew. *Reading and Writing: An Interdisciplinary Journal*, 14, 39–59.
- Badawi, E. M. (1985). Educated spoken Arabic: A problem in teaching Arabic as a foreign language. In K. R. Jankowski (Ed.), *Scientific and humanistic dimensions of language: Festschrift for Robert Lado on the occasion of his 70th birthday on May 31, 1985* (pp. 15–22). Amsterdam: Benjamins.
- Badawi, E.-S. M., Carter, M. G., & Gully, A. (2004). *Modern written Arabic: A comprehensive grammar*. London: Routledge.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Beeston, A. F. L. (1970). *The Arabic language today*. London: Hutchinson.
- Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*, 31, 25–44.
- Buckwalter, T., & Parkinson, D. B. (2011). *A frequency dictionary of Arabic: Core vocabulary for learners*. London: Routledge.
- Carter, R. (1999). Standard grammars, spoken grammars: Some educational implications. In T. Bex & R. J. Watts (Eds.), *Standard English: The widening debate* (pp. 149–166). London: Routledge.
- Chahine, N. (2012). *Reading Arabic: Legibility studies for the Arabic script*. Leiden: Leiden University.
- Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in words and sentences. In R. P. G. Van Gompel, J. H. Fisher, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 341–372). Oxford: Elsevier.
- Dank, M., Deutsch, A., & Bock, K. (2015). Resolving conflicts in natural and grammatical gender agreement: Evidence from eye movements. *Journal of Psycholinguistic Research*, 44, 435–467.
- Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language*, 45, 200–224.
- Ferguson, C. A. (1959). Diglossia. *Word: Journal of the International Linguistic Association*, 15, 325–340.
- Ferguson, C. A. (1996). Epilogue: Diglossia revisited. In A. Elgibali (Ed.), *Understanding Arabic: Essays in contemporary Arabic linguistics in honor of El-Said Badawi* (pp. 49–67). Cairo: American University in Cairo Press.
- Fischer, W., & Jastrow, O. (1980). Einleitung [Introduction]. In W. Fischer & O. Jastrow (Eds.), *Handbuch der arabischen Dialekte [Handbook of the Arabic dialects]* (pp. 15–48). Wiesbaden: Harrassowitz.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178–210.
- Hallberg, A. (2016). *Case endings in spoken Standard Arabic: Statistics, norms, and diversity in unscripted formal speech*. Lund: Lund University.
- Hermena, E. W., Drieghe, D., Hellmuth, S., & Liversedge, S. P. (2015). Processing of Arabic diacritical marks: Phonological/Syntactic disambiguation of homographic verbs and visual crowding effects. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 494–507.
- Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Washington, DC: Georgetown University Press.
- Ibrahim, M. H. (1983). Linguistic distance and literacy in Arabic. *Journal of Pragmatics: An Interdisciplinary Journal of Language Studies*, 7, 507–515.
- Jordan, T. R., Almabruk, A. A. A., Gadalla, E. A., McGowan, V. A., White, S. J., Abedipour, L., et al. (2013). Direction and the central perceptual span: Evidence from Arabic and English. *Psychonomic Bulletin & Review*, 21, 505–511.
- Jordan, T. R., Almabruk, A. A. A., McGowan, V. A., & Paterson, K. B. (2011). Evaluating hemispheric divisions in processing fixated words: The evidence from Arabic. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 47, 992–997.
- Juhász, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1312–1318.
- Juhász, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition*, 13, 846–863.
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773–795.
- Kaye, A. S. (1972). Remarks on diglossia in Arabic: Well-defined vs. Ill-defined. *Linguistics: An Interdisciplinary Journal of the Language Sciences*, 10(81), 32–48.
- Khaldieh, S. A. (2001). The relationship between knowledge of *i'raab*, lexical knowledge, and comprehension of nonnative readers of Arabic. *The Modern Language Journal*, 85, 416–431.
- Kindt, K. T., Højgilt, J., & Kebede, T. A. (2016). Writing change: Diglossia and popular writing practices in Egypt. *Arabica: Journal of Arabic and Islamic Studies*, 63, 324–376.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>.
- Meiseles, G. (1977). Restitution of “word-endings” in Modern Literary Arabic. *Israel Oriental Studies*, 7, 173–195.

- Morey, R. D., Romeijn, J.-W., & Rouder, J. N. (2016). The philosophy of Bayes factors and the quantification of statistical evidence. *Journal of Mathematical Psychology*, 72, 6–18.
- Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of bayes factors for common designs [Computer software]*. Retrieved from <https://CRAN.R-project.org/package=BayesFactor>.
- Nevat, M., Khateb, A., & Prior, A. (2014). When first language is not first: An functional magnetic resonance imaging investigation of the neural basis of diglossia in Arabic. *European Journal of Neuroscience*, 40, 3387–3395. <https://doi.org/10.1111/ejn.12673>.
- Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. (1998). Anomaly detection: Eye movement patterns. *Journal of Psycholinguistic Research*, 27, 515–539.
- Parkinson, D. B. (1993). Knowing Standard Arabic: Testing Egyptians' MSA abilities. In M. Eid & C. Holes (Eds.), *Perspectives on Arabic linguistics V* (pp. 47–73). Amsterdam: Benjamins.
- Parkinson, D. B. (1994). Speaking fuṣḥā. In Y. Suleiman (Ed.), *Arabic sociolinguistics: Issues and perspectives* (pp. 179–211). London: Routledge.
- Paterson, K. B., Almabruk, A. A. A., McGowan, V. A., White, S. J., & Jordan, T. R. (2015). Effects of word length on eye movement control: The evidence from Arabic. *Psychonomic Bulletin & Review*, 22, 1443–1450.
- Pearlmuter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427–456.
- Pickering, M. J., & Garrod, S. (2006). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11, 105–110.
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences: An International Journal of Current Research and Theory with Open Peer Commentary*, 36, 329–347.
- R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Rayner, K. (1998). Eye movements in and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rayner, K. (2009). Eye movements and attention in, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62, 1457–1506.
- Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012). *Psychology of reading*. East Sussex: Psychology Press.
- Rayner, K., & Sereno, S. C. (1994). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory & Cognition*, 22, 281–285.
- Roman, G., & Pavard, B. (1987). A comparative study: How we read in Arabic and French. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements from physiology to cognition* (pp. 431–440). Amsterdam: Elsevier.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16, 225–237.
- Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, 24, 431–451. <https://doi.org/10.1017/S0142176403000225>.
- Saiegh-Haddad, E., & Henkin-Roitfarb, R. (2014). The structure of Arabic language and orthography. In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and perspectives* (pp. 3–28). New York: Springer.
- Saiegh-Haddad, E., & Schiff, R. (2016). The impact of diglossia on vowel and unvoiced word reading in Arabic: A developmental study from childhood to adolescence. *Scientific Studies of Reading*, 20(4), 1–14.
- Trudgill, P. (1975). *Accent, dialect and the school*. London: Arnold.
- Versteegh, C. H. M. (2004). From Classical Arabic to the modern vernaculars. In G. Booij, C. Lehmann, J. Mugdan, & S. Skopetas (Eds.), *Morphology: An international handbook on inflection and word-formation* (Vol. 2, pp. 1740–1754). Berlin: de Gruyter.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.