Sonderdruck aus: / Offprint from:

Zeitschrift für Arabische Linguistik
Journal of Arabic Linguistics
Journal de Linguistique Arabe

Herausgegeben von / Edited by / Edité par
Werner Arnold, Otto Jastrow, Shabo Talay

74 (2021)

Harrassowitz Verlag · Wiesbaden

Manuskripte und Rezensionen werden erbeten an:
Please address manuscripts and book reviews to:

Prof. Dr. Shabo Talay
Freie Universität Berlin
Seminar für Semitistik und Arabistik
Fabeckstrasse 23-25, 14195 Berlin, Germany
Tel. +49 (0)30 838-603 63, Fax +49 (0)30 838-452698
E-Mail: shabo.talay@fu-berlin.de

Diese Zeitschrift kann auch in elektronischer Form über JSTOR (www.jstor.org)
bezogen werden.
This journal can also be accessed electronically via JSTOR (www.jstor.org).

ZAL is a peer reviewed journal with a strict quality control. In a first stage all manuscripts which do not fall within the scope of the journal or do not conform to the requested level of scholarship are sorted out. The remaining manuscripts are then evaluated and selected according to standards of novelty, professionalism, and relevance.

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über https://www.dnb.de/ abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the internet at https://www.dnb.de/.

https://www.harrassowitz-verlag.de/

Satz: Shabo Talay
Gedruckt auf alterungsbeständigem Papier.
Druck und Verarbeitung: Hubert & Co., Göttingen
Printed in Germany

ISSN 0170-026X
eISSN 2747-4437

Inhalt

Aufsätze

Case and mood inflection in Formal Spoken Arabic – a quantitative investigation

Andreas Hallberg[1]

This article presents results from a quantitative investigation of the use of case and mood inflection (CMI) in seventeen televised news-interviews (38,000 words) with Arab politicians and public figures with a variety of backgrounds. The speech of the interviewees was transcribed and annotated for realization of CMI and for relevant morphosyntactic parameters. Speakers were found to vary in the amount of CMI in their speech, but all perform far below the prescriptive ideal, with realizations of case inflection ranging from 0.2% to 42.3% (mean 7.5%, median 3.6%) and mood inflection from 0% to 68.5% (mean 9.9%, median 2.5%). Furthermore, speakers show striking similarities in how the CMI that they do use is distributed in morphosyntactic contexts. First, CMI is for all speakers used at markedly higher rates in words with enclitic pronouns and on words where CMI is orthographically represented. Second, for words with definite article, CMI is almost completely absent. Implications of these results for Arabic language instruction are discussed.

Keywords: case, mood, spoken language, morphology, Standard Arabic, corpus linguistics

0. Introduction

As a standard variety, Standard Arabic (SA) is in the language community considered to be regulated by rules put down in authoritative grammars and lexica. These rules, the codified grammar of SA, have been largely fixed since they were systematized in the 8th to 10th centuries (BOHAS ET AL. 1990), and a gap has arisen between SA as codified and SA as it is used today. The Prague structuralist referred to such a gap as the difference between "codification" and "norm" (HAVRÁNEK 1982). Scholars of Arabic have referred to this with different terms, such as the "ideological" and the "organic standard" (WAHBA 2006) or the "overt" and "covert" norms of SA (PARKINSON 1993). This gap is

1 University of Gothenburg, Department of Languages and Literatures; Göteborg, Sweden. Email: andreas.hallberg@sprak.gu.se

particularity striking for Arabic as a medium for oral communication, a function of SA which has increased in importance since the advent of mass literacy in the Arabic speaking world in the latter half of the 20th century (MAAMOURI 1998) and the expansion of mass media, particularity satellite television (HAMMOND 2007). Speakers in formal public situations are expected to use some form of SA, yet few, if any, are able to comply with the norms of codified grammar in unscripted speech. This has led to the development of linguistic conventions for spoken formal situations that are partly at odds with the codified norm. This article investigates one of the most striking and ideologically loaded differences between the codified norms of SA and the conventions of formal speech, namely the use of case and mood inflection (CMI).[2]

Speech in formal public situations, where speakers are expected to use SA, is characterized by a mixture of standard and non-standard features. Various approaches have been explored to describe this form of Arabic: by dividing the diglossic continuum into "levels", each characterized by a set of linguistic features or by association with particular groups or situations (BLANC 1964; BADAWĪ 1973; MEISELES 1980); by proposing a more or less stable "middle" register (EL-HASSAN 1977; EL-HASSAN 1978); or as code-switching (e.g., BASSIOUNEY 2006; EID 1988; ALBIRINI 2011). Other studies have investigated this variation by quantitatively comparing speakers acting in the same formal, public setting, in radio broadcast discussion programs (SCHULZ 1981), academic panel discussions (MEJDELL 2006), or by instructing informants to use SA in an interview (PARKINSON 1994a; PARKINSON 2003). These latter studies have shown a large degree of individual variation on all variables studied, so that speakers acting in the same situation differ in what SA features they employ and to what extent, and in which non-standard features they avoid. For the purposes of this study, the register under investigation will be referred to as Formal Spoken Arabic, defined as the most formal register available to an educated speaker in unscripted speech. The use of the indefinite in "an educated speaker" is meant to imply that this register may differ between speakers, depending on their background, their experience in speaking in formal public settings, and their views on what linguistic features are appropriate for a given situation.

CMI has a unique position as a marker of standardness and linguistic correctness. CMI, and especially case inflection, has been the focus of Arabic grammar writing since the early stages of codification of Arabic in the eighth century (VERSTEEGH 1983; BOHAS et al. 1990: 50; AYOUB 2006: 629f; YOUNES 2007: 245ff) and continues to this day to be the focal point of Arabic language instruction in Arabic speaking countries (IBRAHIM 1983: 512; MAAMOURI 1998: 53; HAERI 2003: 40; UHLMANN 2012: 104). Yet it is exceedingly rare even for highly educated speakers of Arabic to be able to correctly and consistently

2 Some of these data have been previously presented in HALLBERG (2016).

use inflectional forms in a manner complying with traditional prescriptive grammar (Beeston 1970: 53; Kaye 1972: 43; Ibrahim 1983: 511; Parkinson 1994b; Parkinson 1994a; Saiegh-haddad - Schiff 2016: 3). There are several reasons for this. First, CMI has no parallel in the spoken non-standard varieties and is therefore acquired primarily through formal instruction, and typically through traditional and highly formalized teaching methods based on memorization and repetition, rather than on understanding (Bohas et al. 1990: 16; Uhlmann 2012: 103; Hallberg 2016: 67–71). Second, for syntactic disambiguation and reading comprehension, CMI is almost completely superfluous, since case roles are determined by word order and verbal agreement, and mood by sets of particles preceding the verb (Corriente 1971; Corriente 1973; Holes 2004: 17, 173). Third, forms of CMI that lack an orthographic representation are generally assumed not to be phonologically encoded in silent reading by skilled readers (Bateson 1967: 81f; Stetkevych 2006: 84; Saiegh-haddad - Schiff 2016), and even forms of CMI that are orthographically marked are often not parsed as encoding grammatical information (Hallberg - Niehorster 2021). Speakers of Arabic, accordingly, have little practical incentive to master the systems of CMI, other than for oral performances in formal public settings.

In the context of Arabic tuition at the university level in the West, the question of the appropriate role of CMI in the curriculum is controversial. Arabic teaching materials group rather neatly into two schools. The first school follow the Arabic pedagogical tradition of presenting CMI as a core and mandatory part of proficiency in SA, presenting words in inflected form from the early beginner level (Schulz et al. 2000; Frangieh 2012; Alosh - Clark 2013). The second school ascribe CMI a more peripheral role in SA proficiency, presenting words in uninflected form, and introduce it only at later stages of learning and primarily as passive knowledge (Wightwick - Gaafar 2005; Brustad et al. 2011; Younes et al. 2014). These approaches could both gain from, and be evaluated by, a detailed description of the use of CMI in formal spoken situations.

We currently have only limited knowledge of the extent to which CMI is used in Formal Spoken Arabic, which forms of CMI are typically used, and how this varies between speakers. Meiseles (1977) and Magidow (2012) investigated CMI by analyzing selected example sentences from radio and television broadcasts. This method is problematic due to the variation in the use of CMI even within individual speakers, making it difficult to draw wider conclusions from selected examples. Schulz (1981) investigated occurrences of a number of standard and non-standard features, including CMI, in discussion programs on Egyptian radio. He found a wide variation in the use of CMI. Of the 49 speakers in his corpus, 29 never inflected verbs for mood and 17 speakers never inflected nominals for case. For speakers who did use case inflection it was more common in the accusative, largely due to speakers producing the accusative *-an* ending, which the author attributed to this

ending being orthographically represented. PARKINSON (1994a) provides detailed descriptions of SA oral performances by four speakers in Cairo who were asked to speak SA in an interview. These speakers were hand-picked from a larger pool of interviews to represent the width of variation in the collected material. The speaker with the lowest level of education (one year in high school) added CMI more or less at random, seemingly without understanding their function as grammatical markers. At the other extreme, one speaker used CMI consistently, in a performance that was "utterly unique" in its extreme adherence to prescriptive grammar (PARKINSON 1994a: 207) and which "could only come from someone who is somewhat withdrawn from an engaged role in the more day to day society around him" (PARKINSON 1994a: 209). The other two speakers used CMI to different degrees but with similar patterns of distribution: both used CMI at higher rates on words with enclitic pronouns and where the inflection is orthographically represented, and hardly ever on words with definite article. It is unclear, however, if and how speakers were affected by being instructed to speak SA, rather than responding to a situation that itself calls for a formal register.

We thus have evidence of wide variation in the use of CMI by different speakers, and some indications of how this variation is structured. This study builds on previous studies by using a relatively large material of natural (i.e., not solicited) data, and a detailed coding scheme allowing for a more fine-grained analysis of variation in CMI and comparison between speakers.

The aim of the present article is to provide a characterization of the different ways CMI as used by speakers of Formal Spoken Arabic. In doing so, this article sets out to answer two basic questions:

1. How much does the amount of CMI in Formal Spoken Arabic vary between speakers?
2. Is the use of CMI constrained by morphological or syntactic factors?

The rest of this article is organized as follows. In the first section the method is presented, including a description of the material and the coding scheme. Methodological problems concerning lexicalized forms of CMI and the overlap of standard and non-standard forms are also discussed. In the second section, data on the use of CMI is presented and analyzed according to variation between speakers and correlations with morphosyntactic parameters and orthographic representations. In the fourth section, the patterns of usage of CMI that emerged in the analysis are discussed and interpreted as representing covert linguistic norms in Formal Spoken Arabic. Implications of this interpretation for Arabic pedagogy is discussed, as well as an alternative interpretation of the non-use of CMI as being linguistic errors. A summary and conclusion are presented in the fifth and final section.

1. Method

1.1 Material

The material in this study consists of 17 interviews, each with a different political public figure, broadcast in the program *Liqāʾ al-yawm* [Today's meeting] on the Al Jazeera news channel during 2010 and 2011. The program has the format of a traditional "news interview" (HERITAGE 1985; CLAYMAN - HERITAGE 2002) dealing with current political issues in which the interviewee is involved. This show was chosen for analysis due to the use of unscripted formal speech by a variety of speakers, with substantial material for each speaker, and for its consistent form, topic, and setting.

The interviews in the program follow a fixed format: the interviewer greets the viewers, gives a brief presentation of the interviewee, and then proceeds to ask the guest a series of questions. The topic is exclusively politics, ranging from the general, such the state of democracy in a given country, to the specific, such as confronting the interviewee with a controversial statement made by him or his associates.

An initial pool of 118 programs broadcast during 2010 and 2011 were considered for inclusion. Of these, all interviews with native speakers of Syrian (5), Palestinian (7), and Egyptian Arabic (5) were included in the material. This selection was done in order to limit the number of substrate dialects in the material, so as not to make the analysis overly complex (see below). All the included speakers have a university education and are male. (Only two of all 118 programs broadcasts during the period under consideration featured female interviewees.) Ages range from 50 to 81 years (mean 65.1, $SD = 9.8$). The selection includes some prominent figures, such as Muḥammad Mursī of the Muslim Brotherhood in Egypt, who shortly after this interview was elected president of Egypt; Sallām Fayyāḍ, prime minister of Palestine 2007–2013; and Walīd al-Muʿallim, Syrian foreign minister since 2006. See the appendix for the full list of interviewees and URLs to the published recordings.

Thirteen of the seventeen interviews are available as video where it could be confirmed that the speakers were speaking extemporaneously. The four interviews available only as audio include clear characteristics of extemporaneous speech, such as hesitations and false starts, and there is no indication that answers were read aloud. All included interviews were conducted in a television studio, accept for one, with Riyāḍ al-Asʿad, which was conducted in a private living room.

The seventeen interviews are conducted by thirteen different interviewers, with one appearing in three interviews and two interviewers appearing in two. The interviewers all speak Standard Arabic with limited or no use of dialectal features and with minimal use of CMI.

The length of the programs ranges from 21 to 50 minutes. Counting only speech by the interviewees, the material comprises a total of 5 hours and 22 minutes of continuous speech, or 38,000 words.

1.2 Annotation

Transcripts of the interviews were retrieved from the Al Jazeera news site, transliterated, adapted to the CHAT format (MacWhinney 2000) for ease of analysis, and edited to accurately reflect the recording. Nominals and imperfect verbs were annotated for realization of CMI (inflected, uninflected, ambiguously inflected, incorrectly inflected, inaudible, or not applicable) and for morphological and syntactic parameters determining the prescribed form of CMI. Nominals were annotated for (a) realization of case inflection; (b) syntactic role; (c) head/attribute; (d) definiteness; and (e) declension. Verbs were annotated for (a) realization of mood inflection; (b) syntactic position; (c) person; (d) gender; (e) number; and (f) conjugation. Verbs were annotated for conjugation only to the extent to which mood inflection is affected, giving three declensions: standard (e.g., *yaḏhab* 'go'), hollow (e.g., *yaqūl* 'say'), and "defective" (e.g., *yaʾtī* 'come'). Defective verbs were further divided into three categories depending on the final vowel (*-ā*, *-ī*, or *-ū*). Tokens were also segmented for proclitic, stem, and inflectional suffix (e.g., *wa-yaḏhab-ūn* 'and-3M.go-PL.IND').

Some categories of words and expressions are invariably inflected or uninflected for case and mood by all speakers or are otherwise methodologically problematic. These categories were excluded from annotation but marked up for their respective categories according to the following: (a) titles and forms of address (e.g., *ustaḏ*, appr. 'Mister'); (b) proper nouns; (c) repetitions, corrections, and reformulations; (d) formulaic expressions (e.g., *al-ḥamdu lilāh* 'praise to God'); (e) words with a non-standard stem (e.g., *ḥāga* 'thing') or affix (e.g., the *b*-prefix in *b-fakkir* 'IND-1SG.think'); (f) cardinal numbers above ten; (g) quotations (including the reuse of formulations in the interviewer's question).

Furthermore, adverbs were excluded from all analyses. Many adverbs are derived from adjectives with the suffix *-an*, which is commonly analyzed as an accusative case inflection. It is, however, for several reasons best regarded as a derivational rather than inflectional suffix. First, the suffix changes the part of speech of the word from noun or adjective to adverb. Second, many frequent adverbs formed with *-an*, such as *ǧiddan* 'very' and *ayḍan* 'also', do not occur in the un-derived form (**ǧidd*, **ayḍ*). Third, when adverbs are formed productively, the *-an* suffix is used consistently, as opposed to when the ending indicates a case role in the argument structure of the verb. There are, for example, 44 tokens in the present material of adverbs formed from adjectives with final *-ī/-iyy* (the so-called *nisba* adjectives, e.g., *ǧuzʾiyy-an* 'partly'), and in none of these is the *-an* ending omitted. When this same

ending is used to mark a case relation, such as the object, it is often omitted, as detailed below.

A complication in the annotation of CMI is the overlap between SA word forms that encode mood or case, and forms of the vernacular substrate that do not. The nominal suffix *-īn*, for example, encodes masculine, plural, and non-nominative case in SA, but only masculine and plural in non-standard varieties. It is therefore often impossible to assess whether this suffix on a given word is used as a realization of SA or non-standard morphology, especially since speakers switch between the two systems in unpredictable ways. Such overlapping forms were coded as having ambiguous CMI. The most frequent instances of such ambiguous forms of case inflection are (a) *-a* or *-i* preceding the definite article *l-* on the following word; (b) *-īn* in the sound masculine plural declension; (c) *-ū* in "the five nouns" declension with annexation (e.g., *axū-na* 'our brother'); (d) *-ayn/-ēn* in the dual declension; and (e) the enclitic possessive pronouns *-ak, -ik*, and *-u(h)* (2MS, 2FS, and 3MS). The most frequent forms with ambiguous mood inflection are (a) *-a* or *-i* preceding the definite article *l-* on the following word; (b) *-u* in 3MPL; (c) *-i* in 2FS; (d) the enclitic possessive pronouns *-ak* and *-u(h)* (2MS and 3MS); and (e) *-ø* in the jussive mood.[3] In some situations, the principles for coding a word form as ambiguously inflected differed between speakers, depending on their dialectal substrate. For example, in Egyptian Arabic an epenthetic vowel is inserted after the word stem to break up consonant clusters resulting from the addition of an enclitic pronoun (ABOUL-FETOUH 1969; WOIDICH 2006). In the word *dars-a-na* 'lesson-our', for example, the epenthetic vowel *a* occupies the same place as, and is often indistinguishable from, the case marking vowel. Such tokens were coded as ambiguous, but only for speakers of Egyptian Arabic.

Tokens annotated as ambiguous with regard to CMI (16% of the total data) were excluded from most analysis. This, however, introduces a further complication. Some categories only have forms that are inflected, or ambiguously inflected, but not uninflected. For these categories there are no negative observations of inflection, and any count of CMI after the exclusion of ambiguous forms would necessarily yield 100% realization of CMI, which is clearly problematic. To mitigate this, several analyses were performed on a dataset with these categories excluded. This filtered dataset consists of a

3 Regarding the latter, the null-ending is shared by the non-standard verb form and the SA jussive inflected form. Technically, therefore, all tokens with a null-ending are ambiguous. Coding them as such would, however, render all possible realizations of verbs either inflected or ambiguously inflected, making any analysis of occurrences of CMI on verbs impossible. The choice was therefore made to annotate the null-ending on verbs in the indicative and subjunctive positions as uninflected for mood, rather than as ambiguously inflected.

subset of 13,808 tokens and will in the following be referred to as the *disambiguated dataset*.[4]

Furthermore, words with prescriptively incorrect CMI, words where CMI is not applicable (e.g., nominal stems with final *-ā*), and words with inaudible CMI were excluded in the analyses. These together make up 4.5% of the original data.

1.3 Statistical analysis

Significance tests were performed on generalized regression models on non-aggregated data in *R* (R CORE TEAM 2013) with the *lme4* package (BATES ET AL. 2015) and *p*-values were calculated with the *lmerTest* package (KUZNETSOVA ET AL. 2017). Models were performed with speaker intercept and slope as random effects in a maximal random effects structure (BARR ET AL. 2013), with speaker slope removed as a random effect where models did not converge.

2. Analysis

2.1 Overall rates of inflection

The rate of CMI varies considerably between speakers in the material, even though speakers are acting under very similar circumstances. Rates of CMI for speakers as calculated from the disambiguated dataset are listed in table 1, with speakers ordered from top to bottom by rate of case inflection. The use of case inflection ranges from 0.2% (one token) to 42.2%, clustering around the lower end, with 13 of the 17 speakers performing below the mean of 7.5%. Mood inflection is used at markedly higher rates. One speaker inflects verbs for mood at a rate of 68.5%, three speakers around 25%, and the rest below the mean of 9.9%, while three speakers do not inflect a single word for mood. Speakers who use more of one of the two types of inflection also use more of the other, and they therefore rank in similar ways on both measures ($\rho = .77$, $p < .001$ using Spearman's rank coefficient). Note that the speaker with the highest rates of CMI, Tayzīnī, while being something of an extreme, still makes far from prescriptive use of CMI. Note also that non-standard lexemes were excluded from annotation. A speaker who uses a large proportion of non-standard words can therefore potentially still score high on this measure of CMI, as indeed is the case for Abū Maǧd (see below).

4 The following categories of words were excluded in the disambiguated dataset: for nominals (a) "the five nouns" with annexed noun or pronoun (*abū-k 'father-your'*); (b) sound masculine plural (*miṣrīy-ūn/-īn* 'Egyptians-MPL.NOM/-MPL.ACC/GEN'); (c) dual (*miṣrīy-ān/-ayn* 'Egyptians-DUA.NOM/-DUA.ACC/GEN'); and (d) words with the 2MS, 2FS, or 3MS enclitic pronoun (*-ak, -ik, -u(h)*), and for verbs (a) 2FS (*taḏhabīn*); (b) 2MPL (*taḏhabūn*); (c) 3MPL (*yaḏhabūn*); (d) jussive position (*lam yaḏhab* 'he did not go'); and (e) imperative (*iḏhab* 'go.2MS').

Table 1: Rates of case and mood inflection by speaker

	Case		*Mood*	
	% infl.	*n*	% infl.	*n*
Tayzīnī	42.2	516	68.5	92
Badīʿ	18.3	712	23.8	231
al-Qaddūmī	15.6	475	5.5	109
Abū Majd	10.6	595	26.2	149
Mursī	7.4	539	24.3	136
Kayālī	6.9	635	2.1	141
Fayyāḍ	4.6	819	2.8	144
al-Muʿallim	4.1	441	2.0	102
al-Barādiʿī	3.9	664	3.1	261
ʿAbd al-Qādir	3.9	645	0.7	147
al-Xuḍarī	2.0	842	0.5	214
Hilāl	1.9	646	0.0	69
al-ʾAsʿad	1.7	526	0.8	130
Šallaḥ	1.6	1204	2.5	326
Ġalyūn	1.5	675	5.8	121
ʿArīqāt	0.4	711	0.0	151
al-Miṣrī	0.2	461	0.0	98
Mean	7.5		9.9	
Median	3.9		2.5	
Total		11106		2621

All speakers, except al-Miṣrī (see below), use an unmistakably formal register that can be described as SA with occasional use of CMI, and some admixture of non-standard features. The formal register is most clearly seen in the predominant use of standard variants of frequent words, such as the standard *yurīd* 'want' instead of the non-standard *biddu* (Palestinian, Syrian) or *ʿāyiz* (Egyptian), and in the use of SA phonological forms, such as [q] (*qāl* 'say'), instead of the non-standard [ʾ] or [g] (*ʾāl/gāl*).

Three speakers stand out in the material in using styles that deviate from this overall picture. The first is Tayzīnī, Professor of Philosophy at Damascus University and a distinguished Marxist theoretician. He is an outlier in terms of the very high rates of CMI and complete lack of dialectal features in his speech. He also uses some archaic syntactic structures that are highly stylistically marked in modern SA, for example the negative particle *ʾin* in (1).

It is in the example coupled with consistent use of CMI except preceding the sentence final pause.

(1) *wa-huwa inna mā taqūlūna-hu min mā yaḥduṯ-u*
and-it that what 2MPL.say-it of what 3MS.happens-IND

fī sūrīya in huwa illā natīğat-u muʾāmarat-in xāriğiyya
in Syria not it but result-NOM conspiracy-GEN foreign

'And that is: that which you say about what is happening in Syria is that it is nothing but the result of a foreign conspiracy.'
(Tayzīnī, 20:14)

At the other end of the spectrum is al-Miṣrī, a Palestinian businessman and former minister and member of parliament. In the interview he produces only two instances of case inflection (one of which is a sound masculine plural and therefore not included in table 1) and no instances of mood inflection. His speech is dominated by Palestinian Arabic features, with the occasional admixture of SA, as opposed to the other way around, making him the only speaker in the material who does not use an unmistakably formal register. He uses, for example, the non-standard Palestino-Syrian pseudo-verb *biddu* 'want' 49 times in the interview and the SA variant *yurīd* only twice. The other sixteen speakers together use *biddu* or *ʿāyiz* 30 times. While his speech stands out with a high proportion of non-standard features, his infrequent use of CMI is by no means exceptional: two other speakers make no use of mood inflection and five others inflect words for case at rates below two percent.

Finally, Abū Mağd, Minister of Media in Egypt, stands out in his frequent and drastic switches between saliently standard and non-standard forms. He has a relative high rate of CMI (10.6% case inflection and 26.2% mood inflection), but he combines this with heavy use of saliently non-standard words, often in the same utterance. In (2), for example, the first part of the utterance is saliently non-standard Egyptian Arabic (marked with bold face in the example), and the second part is SA, with standard phonology and with CMI on two of the three words.

(2) ***ana miš ʿāyiz*** *qināʿ yukallim-u qināʿ-an*
I not want mask 3MS.address-IND mask-ACC

'I do not want a mask talking to a mask.' (Abū Mağd, 24:03)

These three speakers illustrate the wide range and variability in the use of CMI in individual styles of Formal Spoken Arabic. Despite this variability, there are some striking similarities in how speakers distribute the CMI that they do use in different morphosyntactic contexts, as shown below.

2.2 Inflection by syntactic role

The most obvious potential pattern of CMI to be tested is one of syntactic position whereby one specific case or mood is inflected more than other positions. This was, however, not the case in this material. With regard to mood, verbs in the disambiguated dataset are inflected for the indicative at a rate of 8.7% ($n = 1995$), and for the subjunctive at 8.9% ($n = 626$). The small difference is not significant. The jussive and the imperative moods do not differ from non-standard forms in terms of inflectional endings and were not included in the disambiguated dataset. (For the special case of jussive inflection of hollow verbs, see below.)

For nominals, words are more often inflected for the accusative than for the nominative or genitive, but this is best explained by orthographic features rather than by syntax. In the disambiguated dataset, 4.8% ($n = 2981$) of nominals in nominative position and 4.7% ($n = 6183$) in genitive position are inflected for case, while 16.1% ($n = 1942$) in accusative position are inflected for case. The much higher rate of inflection in the accusative is significantly different from the rate of inflection in both the nominative ($z = 6.33$, $p < .001$) and the genitive ($z = 7.11$, $p < .001$). Indefinite accusatives in the highly frequent triptote declension are inflected for case with the suffix *-an*, which is orthographically represented with the letter *alif*. If indefinite nouns are excluded, however, the difference in case inflection disappears: 5.85% ($n = 1591$) for nominative, 4.0% ($n = 4428$) for genitive, and 4.7% ($n = 1092$) for accusative, with none of these differences being significant. Further evidence for this being an effect of orthography rather than syntax is presented below.

Overall, then, if the effect of orthography is controlled for, the different cases and moods are inflected at similar rates.

2.3 Words favored for inflection

Two categories of words were found to be highly favored for CMI. First, words with enclitic pronouns show rates of CMI far above the overall average. Rates of CMI in words with and without enclitic pronouns in the disambiguated dataset are listed in table 2 and illustrated in figure 1.

Table 2: Rates of inflection in words with and without enclitic pronoun

	Without enclitic pronoun		With enclitic pronoun	
	% infl.	*n*	% infl.	*n*
Case	5.3	10664	40.5	442
Mood	7.3	2504	41.0	117

Words without enclitic pronoun are inflected occasionally, at 5.3% and 7.3% for case and mood respectively. Words with enclitic pronouns are much less

frequent with only a couple of hundred tokens, or 4% of the total data, but are inflected at very high rates: 40.4% and 41.0% for case and mood respectively. The differences in rates of CMI are significant for both case ($z = 24.45$, $p < .001$) and mood ($z = 9.49$, $p < .001$). The effect is very consistent across speakers, as illustrated in figure 1. Except for a few speakers at the lower end, all speakers drastically increase their rates of CMI in words with enclitic pronouns, as compared to other words.

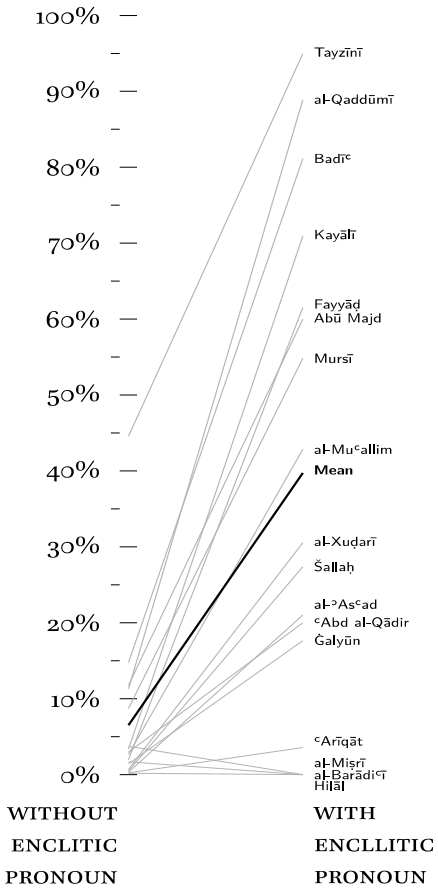

Figure 1: Rates of case and mood inflection in words with and without enclitic pronouns.

The second category of words that have a high rate of CMI in speech are those where the inflectional ending has an orthographic representation. In writing, CMI is mostly absent due to the defective nature of the Arabic writing system, in which short vowels and nunation are represented with optional diacritics. Forms of CMI can thus be categorized as being *orthographic* or *non-*

orthographic, depending on whether the inflectional ending is represented by an addition or change in the letter sequence of the word, or only by an optional diacritic. Orthographic CMI occurs in six situations for case and in four situations for mood, listed in table 3. These situations are highly diverse, representing various combinations of case and mood with inflectional paradigms, definiteness, grammatical person, and properties of the word stem. Because of this diversity, orthographic and non-orthographic CMI can often not be straightforwardly compared without conflation with other variables. Furthermore, some forms of orthographic CMI are very scarce in the data. Nevertheless, the counts of inflected and uninflected tokens with orthographic CMI give a clear indication that tokens with orthographic CMI are inflected at high rates. Types of orthographically marked CMI with 100 or more tokens (A1, A2, and B1 in table 3) all have rates of inflection of 40% and above, to be compared to the overall speaker average of 7.5% and 9.9% for case and mood respectively.

Table 3: Rates of orthographic case and mood inflection. The inflected and uninflected forms in parenthesis are separated by a slash.

		%infl.	*n*
A.	CASE		
1.	Indefinites triptote accusative[5] (*bayt-an/bayt* بيت/بيتا ‘house’)	45.5	497
2.	Sound masculine plural nominative (*miṣriyy-ūn/miṣriyy-īn* مصريين/مصريون ‘Egyptians’)	39.6	101
3.	Triptotes with stem-final ʾ and enclitic pronoun (*mabdaʾ-u-na, mabdaʾ-a-na, mabdaʾ-i-na/mabdaʾ-na* مبدؤنا, مبدأنا, مبدئنا ‘our principle’)	84.2	19
4.	Dual nominative (*dawlat-ān/dawlat-ayn*~*dawlat-ēn* دولتين/دولتان ‘two states’)	41.2	17
5.	“The five nouns” accusative or genitive in annexed form (*ʾax-ī, ʾax-ā/ʾax-ū* أخي, أخا/أخو ‘brother’)	33.3	3
6.	Defective accusative indefinite (*qāḍiy-an/qāḍi* قاضيا/قاض ‘judge’)	0.0	1

5 Excluding stems with final *tāʾ marbūṭa* ـة, *aʾ* ـأ, or *āʾ* ـاء, for which the orthographic CMI is canceled.

B.	MOOD		
1.	2/3MPL indicative	59.2	162
	(*yaḏhab-ūn(a)/yaḏhab-u* يذهبوا/يذهبون ‘go.3MPL’)		
2.	1S, 1PL, 2MS, and 3MS hollow jussive	97.7	43
	(*yakun/yakūn* يكن/يكون ‘was.3MS not’)		
3.	2/3DUA indicative	100.0	2
	(*yaḏhab-ān, yaḏhab-ā* يذهبان, يذهبا ‘go.2MDUA’)		
4.	2FS indicative	n/a	0
	(*taḏhab-īn/taḏhab-i* تذهبين/تذهبي ‘go.2FS’)		

The pattern of high rates of inflection of orthographic CMI is most clearly shown in indefinite accusative nominals of the triptote declension (A1 in table 3), for which there is more data than in the other categories. For indefinite triptotes in nominative and genitive positions, where case inflection is non-orthographic, 4.6% ($n = 2800$) of tokens are inflected for case. For tokens in accusative position, where case inflection is orthographic, 32.9% ($n = 1389$) of tokens are inflected. This difference is significant ($z = 8.28$, $p < .001$). Indeed, every single speaker has a higher rate of case inflection in the accusative than in the other two cases for this type of word. Note that lexicalized forms such as *kaṯīran* ‘a lot’ or *abadan* ‘ever’ are not included in these numbers. However, in doing this comparison, we not only compare orthographic and non-orthographic CMI, but also syntactic role (nominative/genitive vs. accusative), and it is therefore possible that the increased rate of inflection is and effect of the syntactic role, rather than of orthography. To establish the effect of orthography on the rate of case inflection independently of syntactic role, indefinite triptotes were divided into two groups, depending in whether the prescribed accusative ending is orthographic or non-orthographic. This is made possible since the orthographic representation of the accusative marking suffix *-an* is canceled on words stems with the feminine ending *-a(t)* ـة and on nouns with stem-final *-aʾ* ـأ, or *-āʾ* ـاء. Comparing rates of inflection in these two groups of indefinite accusative triptotes isolates the effect of orthography, since the phonological form of the ending and the syntactic role is the same for all words being compared. For words with prescribed orthographic *-an*, the suffix is realized at a rate of 45.5% ($n = 497$), and for words with prescribed non-orthographic *-an*, the corresponding rate is 10.0% ($n = 271$). The difference is significant ($z = 9.4$, $p < .001$). This effect is highly consistent across speakers, as illustrated in figure 2. Thus, in the same morphological and syntactic context, where the accusative inflection is phonologically identical, the rate of inflection in speech is highly affected by the orthographic representation.

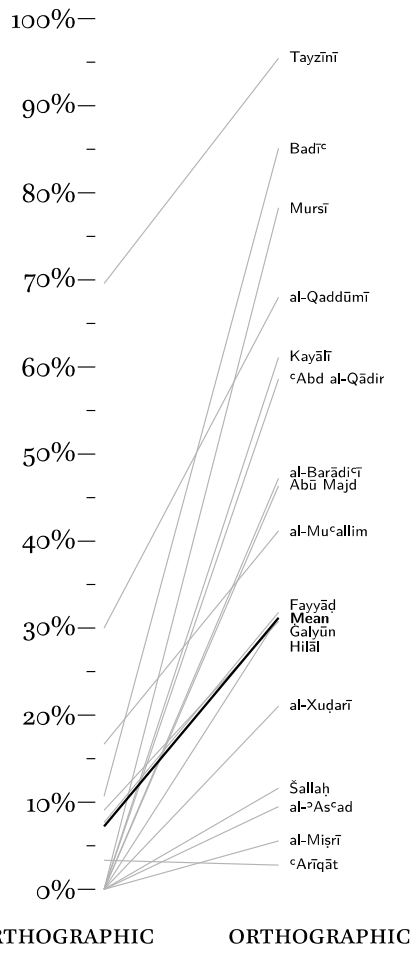

Figure 2: Rates of case inflection for words with a prescribed orthographic and non-orthographic accusative *-an*.

A particularly interesting case of orthographic CMI is inflection in the jussive mood of the so-called hollow verbs (B2 in table 3). This form seems to have a special status in the system of CMI in Formal Spoken Arabic in being applied consistently. In this class of verbs, the middle root surfaces as a stem-internal long vowel, as illustrated in (3). In the jussive mood, which most importantly occurs after the past negative particle *lam*, the long vowel is shortened, as in (4). This is the only form of CMI that affects the internal form of the stem.

(3) *yaqūl(-u)*
3MS.say(-IND)
'he says'

(4) *lam yaqul*
NEG.PAST 3MS.say
'he did not say'

There are 43 instances in the material of hollow verbs in a position where they prescriptively take a shortened jussive form, represented in fourteen of the seventeen speakers. These tokens are inflected for jussive with the shortened form in all instances except one, reproduced in (5). Here, the verb *tasīr* 'proceed' is in jussive position after the negation *lam*, but it is not produced by the speaker in its shortened jussive form *tasir*, as would be prescriptively correct.

(5) *al-qaḍāya l-iqtṣādiyya kull-ha lam tasīr*
DEF-matters DEF-economic all-3FS NEG.PAST 3FS.proceed
bi-manṭiq iqtiṣādi ʿilmi
with-logic economic scientific

'The economic matters did not proceed in a logical, economical, and scientific manner.' (Kayālī, 12:04)

There is a second, synonymous way to negate past tense verbs in SA, with the particle *mā* and a verb in the perfect stem, as in (6). This form of past tense negation, while prescriptively corrects, is also used in non-standard varieties of Arabic and is therefore stylistically less formal, and less frequent in SA (VAN MOL 2003: 264). It does, however, provide speakers with a way to opt out of this form of CMI.

(6) *mā qāl*
NEG.PAST 3MS.say

'he did not say'

Nevertheless, for speakers who do use the jussive form of past tense negation, which is the majority of speakers in the present material, the jussive inflection is applied very consistently. This could not be tested statistically against other forms of jussive inflection, since in other words the jussive inflected form coincides with the non-standard form and is therefore ambiguous. It is nevertheless clear that the shortened jussive form of hollow verbs is used very consistently.

2.4 Words disfavored for inflection

Above, two situations where CMI is highly favored were identified. Variation in CMI may also be constrained in that some forms of CMI are disfavored. Such situations are, however, more difficult to detect due to the low over-all

rates of inflection. For mood, no position that is disfavored for inflection could reliably be identified. One candidate is the subjunctive *-a* inflection on verbs of the defective paradigm (e.g., *yantahiy-a* '3MS.end-SUB' and *yanmuw-a* '3MS.grow-SUB'). There are 43 tokens in the material of defective verbs with prescribed *-a* ending, none of which are inflected for mood, indicating that this inflected form may not be part of Formal Spoken Arabic. This zero-rate of inflection was, however, not significantly different from the 9.6% ($n = 586$) rate of inflection of tokens in other declension with prescribed subjunctive *-a* inflection ($z = .063$, $p = .95$).

For nominals, on the other hand, words with the definite article *(a)l-* stand out as clearly disfavored for case inflection. Nominals were coded for four types of definiteness: (a) definite article; (b) construct state; (c) indefinite; and (d) enclitic pronoun. Examples of these types of definiteness together with percentages of case inflection in the disambiguated dataset are listed in table 4 and illustrated in figure 3. As was shown above, words with enclitic pronouns are strongly favored for case marking. Words with definite articles are, on the other hand, clearly disfavored. Only 1.15% of tokens with the definite articles in the disambiguated dataset are inflected for case, making them the type of definiteness with the lowest rate of inflection. The rate of inflection in words with definite article is different from that of each of the three other types of definiteness ($z > 11$, $p < .001$). Note that tokens with orthographic CMI are excluded from this dataset, so that the pattern of words with definite articles being uninflected for case only holds for non-orthographic case inflection. Words with definite article make up around half of all nominals in the material, meaning that the fact that these are disfavored for case marking in and of itself accounts for a large part of the overall lack of CMI.

Table 4: Percentage of case inflection by type of definiteness

	% infl.	*n*	*Example*	
Definite article	1.1	5136	*al-bayt(-u)*	DEF-house(-NOM)
Construct state	5.4	1533	*bayt(-u)*	house(-NOM)
			aḥmad	Ahmed
Indefinite	11.7	3994	*bayt(-un)*	house(-NOM)
Enclitic pronoun	40.5	441	*bayt(-u)-na*	house(-NOM)-our

This pattern is consistent across speakers, as illustrated in figure 3. Tayzīnī, the speaker with by far the highest rate of CMI, stands out in inflecting words with definite articles at the very high a rate of 10.8%. In fact, of the 59 tokens of inflected case on words with the definite articles in the disambiguated dataset, around half, 29 tokens, are produced by this one speaker. This high

rate is far below his rates in other types of definiteness, however, indicating that his use of CMI is constrained by similar overall patterns as for other speakers.

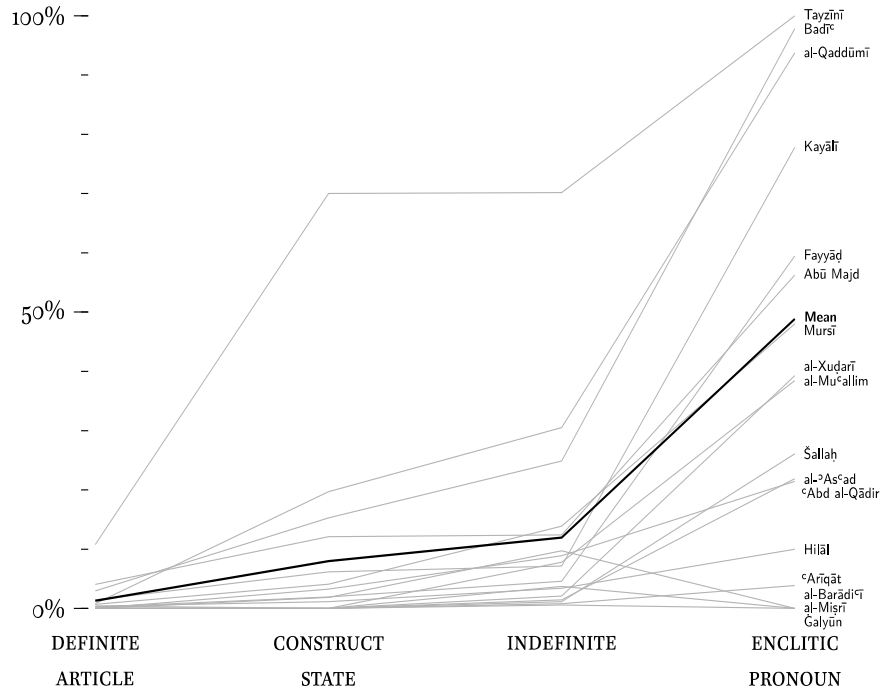

Figure 3: Rates of case inflection by definiteness.

3. Discussion

This article set out to answer two questions about the variation in CMI in Formal Spoken Arabic: how much it varies between speakers and whether it is constrained by morphological or syntactic factors. Regarding the first question, this study aligns with previous research in finding large measure of variation between speakers in the amount of CMI used. It adds to this research by providing a quantification of this variation. Speakers inflect on average 7.5% of nominals for case (median 3.9%) and 9.9% of verbs for mood (median 2.5%). The much lower medians indicate that the distribution is highly skewed towards the lower end, with rates above the mean quickly becoming increasingly rare.

Regarding the question of how the variation in CMI is constrained by morphological and syntactic factors, three distinct patterns emerged. First, non-orthographic CMI is very rare in nominals with definite articles. This has far reaching consequences, since these words constitute a large part of natural speech: 65.6% of all nominal tokens, and 53.2% of the total number of tokens (nominal and verbal) are of this category. This pattern by itself, therefore, sets a ceiling level for the global rate of CMI below 50%. This is especially noteworthy, since, as a consistent pattern of *non*-use of CMI, it is in direct conflict with prescriptive grammar.

The second way in which the use of CMI is structured in Formal Spoken Arabic is that all speakers, to the extent that they employ CMI, do so at much higher rates in words with orthographic CMI. The preference for using orthographic CMI is most likely an effect of reading. Given the fact that Formal Spoken Arabic is a highly specialized register with low density of communication providing limited aural input, we may assume that speakers draw heavily on written input for developing proficiency in Formal Spoken Arabic. In written SA, CMI is only represented in a few limited situations, but with complete consistency. These forms are probably, therefore, more effectively internalized by speakers than are other forms, and speakers may draw on this proficiency for stylistic variation by means of CMI in their speech, resulting in a disproportionate amount of orthographic CMI.

The third pattern is that speakers produce high rates of CMI on words with enclitic pronouns. A possible explanation for this is that it is the result of reanalysis of aural input. One of the sources of input for developing proficiency in CMI may be prescriptively correct renderings of SA, as heard in news broadcasts, voice-overs, and recitation. In these forms of SA, omission of diacritic CMI is mandatory preceding a breath pause (MITCHELL 1990: 99; NELSON 2001: 29; HOLES 2004: 62). Forms such as *balad-ø* '[a] country' or *al-balad-ø* 'DEF-country-ø', without CMI, are therefore common in these forms of Arabic, and furthermore occur in perceptually salient positions preceding a pause at the end of clauses or major clause constituents. Speakers may draw on such performances and generalize the pausal omission of CMI to also include other, non-pausal positions, giving, for example, the consistent omission of CMI in words with definite articles demonstrated above. (It is unclear, however, why this would affect words with definite article more than indefinite words.) The prescriptive omission of CMI is not applied to words with enclitic pronouns, so that forms such as *balad-na* 'country-our', with an enclitic pronoun and the omission of CMI, never occur in prescriptively correct renderings, potentially blocking the generalization of CMI omission in these words.

Whatever the explanations may be for how these patterns have arisen, they appear, on the basis of this data, to have been established as covert conventions for Formal Spoken Arabic. That is, in their use of CMI, speakers adhere to the following: (a) to the extent that CMI is added, it is added

primarily to words with enclitic pronouns and on words where the CMI is orthographically represented, and (b) it is not, or extremely rarely, added to words with definite article. Put differently, certain positions are open for free variation in CMI, whereas other positions are closed for such variation and excluded from CMI.

The interpretation of these patterns as reflecting covert conventions predicts that deviations from these patterns are unconventional, and would accordingly be negatively evaluated by listeners (cf. PARKINSON 1994a). An example of such an unconventional utterance is (7), adapted from (5). In this sentence, CMI is added on words with definite article, but not elsewhere, giving a sentence that deviates from the covert conventions described above, while complying with overt norms, in that the CMI that is added is prescriptively correct.

(7) *al-qaḍāyat-u* *l-iqtṣādiyyat-u* *kull-ha* *lam*
DEF-matters-NOM DEF-economic-NOM all-3FS NEG.PAST
tasir *bi-manṭiq* *iqtiṣādi* *ʿilmi*
3FS.proceed with-logic economic scientific

'The economic matters did not proceed in a logical, economical, and scientific manner.'

Evidence that too much use of CMI is negatively evaluated is presented in PARKINSON (1991). In a matched-guise experimental design, participants listen to different forms of reading aloud of the same text. The readings differed in the realizations of phonological features and with CMI ranging from complete to none. Participants rated the reader using SA phonology but only *partial* CMI as being more "smart" than the others. PARKINSON (1991: 59) speculates that the lower ranking of the text with full CMI may be due to "the feeling that it is overdone, ponderous, showing off more than necessary, etc." Future research could further test the prediction presented here by presenting participants with discourse with the same amount of CMI, but where it is used in "conventional" and "unconventional" positions according to the patterns described above.

The findings presented in this study have important implications for Arabic language instruction, provided that proficiency in Formal Spoken Arabic (or oral proficiency in SA) is a learner aim. A first implication is that speakers use uninflected forms as the default, to which CMI is occasionally added, as indicated by the low over-all rates of CMI. Accordingly, CMI in Formal Spoken Arabic is best regarded as added through a process of addition, rather than uninflected forms being the result of a process of omission. This suggests that the uninflected forms are best learned before the inflected forms. Training students in using inflected forms, particularity in words with definite article, risks instilling habits that must later to be unlearned in order to develop skills in Formal Spoken Arabic. These default uninflected forms

include the *-īn* suffix for the sound masculine plural, also in nominative position (see table 3 A2) and uninflected forms of words with enclitic pronouns (e.g., *nunāqiš-hā* 'discuss.1PL-it', see table 4).

A second implication for language instruction is that not all forms of CMI are equal. Orthographic CMI is, clearly, important in correct writing (although less so for reading, HALLBERG – NIEHORSTER 2021), and, as this study has demonstrated, is also heavily used for stylistic variation in speech. Non-orthographic CMI, on the other hand, is peripheral to writing and in reading comprehension of most texts, since it is not present in normal undiacritized writing, and it is not freely used for stylistic variation in formal speech. Proficiency in non-orthographic CMI is therefore for most learners of marginal benefit.

An alternative interpretation of the patterns of (non-)use of CMI presented above is that they represent linguistic errors, and that speakers, for example those in the present material aim at, but fail to reach, complete use of CMI. This view is worth commenting on, seeing to its prevalence in the Arabic language community at large as well as among Arabic teachers and specialists, as reflected, for example, in commonly used teaching materials. According to this view, the target form for learners developing skills in a formal spoken register ought to be prescriptively correct speech with consistent use of CMI, as described in traditional grammars. The way speakers in this study use CMI should therefore not be taken as a reference-point for learners. Rather, their way of using CMI is seen as an example of the deplorable state of SA proficiency even among the educated elite. This view aligns with standard language ideology, characterized by "intolerance of optional variability in language" (MILROY – MILROY 1991: 26), in this context variability in CMI, and attempts to hinder, or even revert, real or perceived linguistic change (MILROY 1999; RICENTO 2006). Furthermore, this view of CMI in speech can be seen as an expression of what SULEIMAN (1996) has called the "naïve-realist" conception in the Arabic grammatical tradition, according to which the grammatical description (as formulated in traditional Arabic grammars) is taken to be logically identical to "Arabic in its pre-descriptive state". In this conception of SA, only the traditional codified norm has legitimacy, and observations of language in-use are seen as inconsequential for what is to be considered appropriate use of the language.

Having prescriptively correct use of CMI as the target form for proficiency training in the formal register, accordingly, does not prepare students to become integrated into the community of users of this register, but trains them to speak in way that deviates from and challenges the speech habits of that community. This may not be in the interest of the learner.

4. Summary and conclusion

This study investigated the use of CMI in Formal Spoken Arabic, represented by televised news-interviews with public figures who are native speakers of Arabic. The study found a wide individual variation in rates of CMI, but with all speakers performing far below the prescriptive ideal. Within this variation there were striking similarities in how the CMI that they did use was distributed by morphosyntactic parameters. On the one hand, speakers used CMI at higher rates on words where the inflectional ending has an orthographic representation, as well as on words with enclitic pronoun. On the other hand, CMI was extremely rare on words with definite article. It was argued that these patterns of variation in CMI reflect covert linguistic norms in Formal Spoken Arabic, and that these norms should be reflected in the formulation of proficiency aims and curriculum design for Arabic language instruction.

5. Appendix

Below are listed the names, nationalities, and occupations or positions of speakers in the material, as well as URLs to the recorded interviews. Interviews are available either as video on Al Jazeera's YouTube channel, or as audio on their website.[6]

1. Ḥātim ʿAbd al-Qādir. Palestine. Minister of Jerusalem Affairs. *https://www.youtube.com/watch?v=kR4cQZjZPQU*
2. Aḥmad Abū l-Majd. Egypt. Jurist; liberal Islamist thinker; former Minister of Youth and Minister of media. *https://www.youtube.com/watch?v=qyoL92NpSN8*
3. Ṣāʾib ʿArīqat. Palestine. Minister of negotiations. *http://www.aljazeera.net/audioplayer/42ef7b37-9c74-49ff-8368-ef4c6b41bcb2*
4. Riyāḍ al-Asʿad. Syria. Founder and commander of the Free Syrian Army; former general, Syrian army. *https://www.youtube.com/watch?v=8arRIIt8PTU*
5. Muḥammad Badīʿ. Egypt. Supreme guide of the Muslim Brotherhood. *https://www.youtube.com/watch?v=qlUOAXUnOFI*
6. Muḥammad al-Barādiʿī. Egypt. Founder of National Association for Change; Nobel Peace Price laureate, former director general of IAEA. *https://www.youtube.com/watch?v=Phi-NPJXUK4*

6 Last confirmed access 27 September 2019.

7. Salām Fayyāḍ. Palestine. Prime Minister.
   http://www.aljazeera.net/audioplayer/a0f24635-4a55-471a-968f-f096bdc2ed2a
8. Burhān Ġalyūn. Syria. Head of the Syrian National Council.
   https://www.youtube.com/watch?v=zkQxo_ejsJc
9. ʿAlī d-Dīn Hilāl. Egypt. Media Secretary, Naitional Democratic Party; former Minister of Youth.
   https://www.youtube.com/watch?v=ht4j0rQrzYU
10. Munīb al-Miṣrī. Palestine. Businessman; former Minister of Public Works, Jordan; former MP, Fath.
    https://www.youtube.com/watch?v=imWxttjk3u8
11. Muḥammad Kayālī. Syria. Former head of military police.
    https://www.youtube.com/watch?v=XvCWHsVmHgk
12. Walīd al-Muʿallim. Syria. Minister of foreign affairs.
    http://www.aljazeera.net/audioplayer/ff65db75-e6d5-4b72-9d54-f5431bac5de9
13. Muḥammad Mursī. Egypt. Chairman of the Freedom and Justice Party; affiliated with the Muslim Brotherhood.
    https://www.youtube.com/watch?v=FJCZ7pR1Img
14. Farūq al-Qaddūmī a.k.a. Abū Luṭf. Palestine. Cofounder of Fath.
    http://www.aljazeera.net/audioplayer/0bfd7ece-f358-4c64-ba68-e6fc7887d157
15. Ramaḍān Šallaḥ. Palestine. General secretary of Islamic Jihad.
    https://www.youtube.com/watch?v=aDJ3_GO13jg
16. Ṭayyib Tayzīnī. Syria. Professor of Philosophy, Damascus University, Marxist thinker.
    https://www.youtube.com/watch?v=tf3XoeGpBr8
17. Jamāl al-Xuḍarī. Palestine. Independent MP; Head of the Popular Committee Against Siege of Gaza.
    https://www.youtube.com/watch?v=rZcd7vAx-hY

6. References

ABOUL-FETOUH, H.M. (1969): A Morphological Study of Egyptian Colloquial Arabic. The Hague: Mouton.

ALBIRINI, A. (2011): The Sociolinguistic Functions of Codeswitching Between Standard Arabic and Dialectal Arabic. In: Language in Society 40(5), 537–562.

ALOSH, M. - CLARK, A.S. (2013): Ahlan Wa Sahlan: Functional Modern Standard Arabic for Intermediate Learners. New Haven; London: Yale University Press.

AYOUB, G. (2006): Laḥn. In: VERSTEEGH, C.H.M. (ed.): Encyclopedia of Arabic language and linguistics, vol. ii. Leiden: Brill, 628–634.

BADAWĪ, A.-S.M. (1973): Mustawayāt al-ʿarabiyya al-muʿāṣira fī miṣr. Cairo: Dār al-maʿārif bi-Miṣr.

BARR, D.J. - LEVY, R. - SCHEEPERS, C. - TILY, H.J. (2013): Random Effects Structure for Confirmatory Hypothesis Testing: Keep It Maximal. In: Journal of Memory and Language 68(3), 255–278.

BASSIOUNEY, R. (2006): Functions of Code Switching in Egypt: Evidence from Monologues. Leiden: Brill.

BATES, D. - MÄCHLER, M. - BOLKER, B. - WALKER, S. (2015): Fitting Linear Mixed-Effects Models Using lme4. In: Journal of Statistical Software 67(1), 1–48.

BATESON, M.C. (1967): Arabic Language Handbook. Washington, D.C: Center for Applied Linguistics.

BEESTON, A.F.L. (1970): The Arabic Language Today. London: Hutchinson.

BLANC, H. (1964): Stylistic Variation in Spoken Arabic: A Sample of Interdialectal Educated Conversation. In: FERGUSON, C.A. (ed.): Contributions to Arabic linguistics. Cambridge, Mass.: Harvard University Press, 79–156.

BOHAS, G. - GUILLAUME, J.-P. - KOULOUGHLI, D.E. (1990): The Arabic Linguistic Tradition. London: Routledge.

BRUSTAD, K. - AL-BATAL, M. - AL-TUNSI, A. (2011): Al-Kitaab Fii Taʿallum Al-ʿArabiyya: A Textbook for Beginning Arabic. Washington, DC: Georgetown University Press.

CLAYMAN, S. - HERITAGE, J. (2002): The News Interview: Journalists and Public Figures on the Air. Cambridge: Cambridge University Press.

CORRIENTE, F. (1973): Again on the Functional Yield of Some Synthetic Devices in Arabic and Semitic Morphology (a Reply to J. Blau). In: The Jewish Quarterly Review 64(2), 154–163.

CORRIENTE, F. (1971): On the Functional Yield of Some Synthetic Devices in Arabic and Semitic Morphology. In: The Jewish Quarterly Review 62(1), 20–50.EID, M. (1988): Principles for Code-Switching Between Standard and Egyptian Arabic. In: Al-Arabiyya 21(1), 51–79.

EL-HASSAN, S.A. (1977): Educated Spoken Arabic in Egypt and the Levant: A Critical Review of Diglossia and Related Concepts. In: Archivum Linguisticum 8, 112–132.

EL-HASSAN, S.A. (1978): Variation in the Demonstrative System in Educated Spoken Arabic. In: Archivum Linguisticum 9, 32–57.

FRANGIEH, B.K. (2012): Arabic for Life: A Textbook for Beginning Arabic. New Haven, Conn.: Yale University Press.

HAERI, N. (2003): Sacred Language, Ordinary People: Dilemmas of Culture and Politics in Egypt. New York: Palgrave Macmillan.

HALLBERG, A. (2016): Case Endings in Spoken Standard Arabic: Statistics, Norms, and Diversity in Unscripted Formal Speech. Lund: Lund University.

HALLBERG, A. - NIEHORSTER, D.C. (2021): Parsing Written Language with Non-Standard Grammar: An Eye-Tracking Study of Case Marking in Arabic. In: Reading and Writing: An Interdisciplinary Journal 34, 27–48.

HAMMOND, A. (2007): Popular Culture in the Arab World: Arts, Politics, and the Media. Cairo: American University in Cairo Press.

HAVRÁNEK, B. (1982): Zum Problem der Norm in der heutigen Sprachwissenschaft und Sprachkultur. In: STEGER, H. (ed.): Soziolinguistik: Ansätze zur soziolinguistischen Theoriebildung. Darmstadt: Wissenschaftliche Buchgesellschaft, 289–296.

HERITAGE, J. (1985): Analyzing News Interviews: Aspects of the Production of Talk for an Overhearing Audience. In: DIJK, T.A. VAN (ed.): Handbook of discourse analysis. London: Academic Press, 95–117.

HOLES, C. (2004): Modern Arabic: Structures, Functions, and Varieties. Washington, DC: Georgetown University Press.
IBRAHIM, M.H. (1983): Linguistic Distance and Literacy in Arabic. In: Journal of Pragmatics 7(5), 507–515.
KAYE, A.S. (1972): Remarks on Diglossia in Arabic: Well-Defined vs. Ill-Defined. In: Linguistics 10(81), 32–48.
KUZNETSOVA, A. - BROCKHOFF, P.B. - CHRISTENSEN, R.H.B. (2017): lmerTest Package: Tests in Linear Mixed Effects Models. In: Journal of Statistical Software 82(13), 1–26.
MAAMOURI, M. (1998): Language Education and Human Development: Arabic Diglossia and Its Impact on the Quality of Education in the Arab Region. Philadelphia: International Literacy Institute.
MACWHINNEY, B. (2000): The CHILDES Project: Tools for Analyzing Talk. Mahwah, NJ: Lawrence Erlbaum Associates.
MAGIDOW, A. (2012): Explaining Inconsistent Case Marking in Spoken Formal Arabic. In: Zeitschrift für arabische Linguistik (55), 62–99.
MEISELES, G. (1980): Educated Spoken Arabic and the Arabic Language Continuum. In: Archivum Linguisticum 11, 118–148.
MEISELES, G. (1977): Restitution of "Word-Endings" in Modern Literary Arabic. In: Israel Oriental Studies 7, 173–195.
MEJDELL, G. (2006): Mixed Styles in Spoken Arabic in Egypt: Somewhere Between Order and Chaos. Boston, MA: Brill.
MILROY, J. (1999): The Consequences of Standardisation in Descriptive Linguistics. In: WATTS, R.J. - BEX, T. (eds.): Standard English: The widening debate. London: Routledge, 16–39.
MILROY, J. - MILROY, L. (1991): Authority in Language: Investigating Language Prescription and Standardisation. London: Routledge.
MITCHELL, T.F. (1990): Pronouncing Arabic. Oxford: Clarendon.
NELSON, K. (2001): The Art of Reciting the Qurʾan. Cairo; New York: The American University in Cairo Press.
PARKINSON, D.B. (1993): Knowing Standard Arabic: Testing Egyptians' MSA Abilities. In: EID, M. - HOLES, C. (eds.): Perspectives on Arabic Linguistics (5). Amsterdam: Benjamins, 47–73.
PARKINSON, D.B. (1991): Searching for Modern Fuṣḥā: Real-Life Formal Arabic. In: Al-Arabiyya 24, 31–64.
PARKINSON, D.B. (1994a): Speaking Fuṣḥā in Cairo. In: SULEIMAN, Y. (ed.): Arabic sociolinguistics: Issues and perspectives. London; New York: Routledge, 179–211.
PARKINSON, D.B. (1994b): Testing Native Speakers: Implications for Teaching Arabic to Non-Native Speakers. In: MCCARUS, E.N. - RAMMUNY, R.M. - PARKINSON, D.B. (eds.): Investigating Arabic: Linguistic, pedagogic and literary studies in honor of Ernest N. McCarus. Columbus, Ohio: Greyden Press, 191–203.
PARKINSON, D.B. (2003): Verbal Features in Oral Fuṣḥa Performances. In: International Journal of the Sociology of Language 163(1), 27–41.
R CORE TEAM (2013): R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.
RICENTO, T. (2006): Language Policy: Theory and Practice: an introduction. In: RICENTO, T. (ed.): An introduction to language policy: Theory and method. Malden, MA: Blackwell Publ, 10–23.

SAIEGH-HADDAD, E. - SCHIFF, R. (2016): The Impact of Diglossia on Voweled and Unvoweled Word Reading in Arabic: A Developmental Study from Childhood to Adolescence. In: Scientific Studies of Reading 20(4), 1–14.
SCHULZ, D.E. (1981): Diglossia and Variation in Formal Spoken Arabic in Egypt. Ann Arbor: Unpublished Ph.D. thesis.
SCHULZ, E. - KRAHL, G. - REUSCHEL, W. (2000): Standard Arabic: An Elementary/ Intermediate Course. Cambridge: Cambridge University Press.
STETKEVYCH, J. (2006): The Modern Arabic Literary Language: Lexical and Stylistic Developments. Washington DC: Georgetown University Press.
SULEIMAN, Y. (1996): The Simplification of Arabic Grammar and the Problematic Nature of the Sources. In: Journal of Semitic Studies 41(1), 99–119.
UHLMANN, A.J. (2012): Arabs and Arabic Grammar Instruction in Israeli Universities: Alterity, Alienation and Dislocation. In: Middle East Critique 21(1), 101–116.
VAN MOL, M. (2003): Variation in Modern Standard Arabic in Radio News Broadcasts: A Synchronic Descriptive Investigation into the Use of Complementary Particles. Leuven; Dudley, Mass: Peeters; Departement Oostere Studies.
VERSTEEGH, K. (1983): Arabic Grammar and Corruption of Speech. In: Al-Abhath 31, 139–60.
WAHBA, K.M. (2006): "First Language Teaching." In: Encyclopedia of Arabic language and linguistics, 2, 104–13.
WIGHTWICK, J. - GAAFAR, M. (2005): Mastering Arabic Grammar. Basingstoke: Palgrave Macmillan.
WOIDICH, M. (2006): Cairo Arabic. In: VERSTEEGH, C.H.M. (ed.): Encyclopedia of Arabic language and linguistics. Leiden: Brill, 323–333.
YOUNES, M. (2007): On 'I"rāb", Power, and Language Reform in the Arab World. In: al-ʿArabiyya 40/41, 221–241.
YOUNES, M.A. - WEATHERSPOON, M. - FOSTER, M.S. (2014): Arabiyyat Al-Naas (Part One) (2 vols.). Abingdon, Oxon; N.Y., NY: Routledge.