

## Chapter 28

# Contact-induced grammaticalization between Arabic dialects

Thomas Leddy-Cecere

Bennington College

This chapter describes the phenomenon of contact-induced grammaticalization between Arabic dialects and its significance in accounting for the development of future tense markers across modern Arabic varieties. After an introduction to theoretical aspects of general grammaticalization theory and contact-induced grammaticalization in particular, discussion shifts to the identification of specific contact-induced grammaticalization processes leading to the modern distribution of future tense-marking forms across the Arabic-speaking world. Finally, the significance of these findings to broader inquiry in Arabic dialectology and theoretical contact linguistics is considered.

## 1 Introduction & theory

### 1.1 Overview

This chapter presents evidence for the occurrence of contact-induced grammaticalization processes between dialects of Arabic over the course of the language's history. The critical role of dialect contact as a source of synchronic variation and diachronic change across Arabic varieties is well recognized, and the description of its outcomes a long-standing occupation of Arabic dialectologists (e.g. Behnstedt & Woidich 2005; Miller et al. 2007). Representing a fairly recent theoretical development in the field of contact linguistics – following largely from the proposals of Heine & Kuteva (2003; 2005) and Dahl (2001) – contact-induced grammaticalization as a model has not been applied to the analysis of Arabic dialect data on a large scale. As will be seen, however, the phenomenon displays significant merit as an account for the evolution and diffusion of a number of morphosyntactic features across the modern Arabic dialects.



In the following subsections, I begin with a review of the current state of research in grammaticalization theory and in contact-induced grammaticalization (CIG) specifically. I then proceed to an illustrative example of CIG in the Arabic context, demonstrating the model's power as an explanatory mechanism in interpreting the distribution of future tense markers across the modern dialects. I conclude the chapter with a brief discussion of the broader significance of CIG in the analysis of Arabic and the potential role for Arabic data in advancing general theoretical knowledge of the phenomenon at large.

## **1.2 Grammaticalization**

Most linguists agree that it is possible to synchronically classify the majority of linguistic forms along a cline from "more lexical" to "more grammatical", in a manner roughly consistent with the progression as conceived by Hopper & Traugott (2003):

CONTENT WORD > GRAMMATICAL WORD > CLITIC > INFLECTIONAL AFFIX

Historical linguists would add to this synchronic observation the diachronic reflection that it is common to observe a single etymological item advancing through the successive stages of this cline as it develops as part of a linguistic system over time. In fact, the sheer frequency of examples indicating such a trajectory of evolution has led to the identification of a cross-linguistically attested phenomenon known as grammaticalization. The following definition provided by Hopper and Traugott is indicative of several currently referenced in the field, which – though differing in emphasis and points of detail – broadly subscribe to a similar central principle:

[Grammaticalization is] the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions and, once grammaticalized, continue to develop new grammatical functions (Hopper & Traugott 2003: 18).

Though useful for purposes of general definition, this largely intuitive formulation of grammaticalization and the cline which it follows must be further deconstructed if they are to be operationalized as part of a rigorous analysis. Andersen (2008) summarizes the issue succinctly in observing that the grammaticalization cline so articulated conflates numerous discrete dimensions of language change by presenting them as unified steps in a chain: the shift from lexical to grammatical word is one of semantic content, while that from word to clitic to affix

involves morphosyntax and any associated loss of phonological material is best understood as phonological change. Since the early stages of grammaticalization research, more complex approaches to the description of the phenomenon have been proposed based on the concurrent evaluation of multiple parameters (e.g. Lehmann 1985). Other authors opt instead to define analogous parameters in terms of diachronic processes, thereby rendering them more directly relatable to the modes of historical linguistic analysis which underlie the bulk of investigations in grammaticalization research. The latter approach is adopted here, largely following the account proposed by Heine (2004).

Heine views grammaticalization as defined by the simultaneous progression of four distinct but interrelated diachronic processes: desemanticization, extension, decategorialization, and erosion. Desemanticization involves the loss of concrete lexical (“content”) meaning and a corresponding rise in abstract grammatical function associated with the use of an item in particular contexts. This often represents the first observable stage of grammaticalizing change, and, as its name suggests, primarily concerns the semantic content of the item rather than its distribution, form, or syntactic behavior. Closely coupled with desemanticization is extension, namely the novel use of the grammaticalizing item in contexts in which it was not previously employed; extension is thus defined as a change in incidence. The hand-in-hand advance of these two processes is demonstrated in the evolution of the French negative element *pas*: having shed its content semantics as a noun meaning ‘step’ and developed grammatical function as a marker of verbal negation, *pas* is extended in contemporary usage to contexts involving none of the implied motion of its lexical source (Hansen & Visconti 2009: 137–138).

The third process described by Heine, that of decategorialization, consists of the changes by which a grammaticalized item comes to lose the morphosyntactic properties characteristic of its source’s original word class, such as word order freedom or agreement inflection; an example may be found in the gradual development of the English adverbial marker *-ly* from a morphosyntactically free substantive meaning ‘body, form’ to a bound derivational suffix (Ramat 2011: 505). Erosion, the fourth process considered by Heine, refers to the gradual reduction and lenition of phonological form beyond what is accounted for by regular sound change, as observed in the irregular changes deriving the Jewish Babylonian Aramaic continuous aspect marker *qā ~ kā* from earlier \*qā?ē ‘standing’ (Rubin 2005: 134).

Theories of grammaticalization have also been strongly linked to the notion of unidirectionality, the proposal that change along the above-described cline occurs only from more lexical to more grammatical and not vice versa (Lehmann

2015). Though the absolute formulation of this hypothesis has been the subject of much debate (e.g. Norde 2009), recognition of a strong unidirectional tendency remains integral to understandings of grammaticalization on both empirical and theoretical grounds (Haspelmath 1998; Heine 2004). It has been proposed that the impetus for such a tendency lies in a universal set of cognitive and communicative principles common to the human mental faculty (Claudi & Heine 1986; Bybee 2003; Lehmann 2015); these would provide an account for the pervasive occurrence of grammaticalization as a worldwide phenomenon, and may be seen to bias the results of grammatical change in the directions entailed by the four processes described above.

The concomitant advancement of these processes is discernible in one of the few cases of Arabic grammaticalization for which a reasonably complete chain of historical development is attested: that of the Egyptian Arabic future tense marker *ḥa-*. Documented in sixteenth- and seventeenth-century sources as *rāyih*, this item already shows evidence of desemanticization and extension, departing from the semantics of its lexical source as an active participle ‘going’ to indicate intention and imminent futurity of action, consequently allowing its extension to usage contexts devoid of actual motion: *ʔanā rāyih aḡannī ʔalēh* ‘I am going to sing about it [and proceeds to sing]’ (Davies 1981: 241). In its nineteenth-century incarnation *rāḥ* ~ *raḥ* ~ *ḥa-*, the form shows further desemanticization and extension as its value changes from an imminent to a general future and it comes to be employed in previously unacceptable circumstances, such as in the presence of a non-immediate temporal adverb: *rāḥ yīgi bukra* ‘he’ll come tomorrow’ (Elias & Elias 1981: 157; for earlier usage constraints, see Davies 1981: 241-243). These increasingly modern forms also attest decategorialization, as the once-obligatory adjectival agreement marking of the participial original becomes optional – *raḥ* (SG.M)/*raḥa* (SG.F)/*raḥīn* (PL) ~ *raḥ* (invar.) (Vollers 1895: 40) – and eventually ceases to exist altogether in the tightly bound modern clitic *ḥa-* ~ *ha-* (Abdel-Massih et al. 2009: 268). Fourthly, progressive phonetic erosion is visible throughout the item’s history, as none of the loss or lenition of phonetic material through the stages *rāyih* > *rāḥ* > *raḥ* > *ḥa-* > *ha-* attested above is attributable to regular sound change. Taken together, these combined processes chart the grammaticalization of lexical *rāyih* ‘going’ through its gradual development into the modern future tense clitic *ḥa-*.<sup>1</sup>

---

<sup>1</sup>On sources referenced in the preceding paragraph: Davies (1981) is a study of colloquial elements in the seventeenth century Egyptian text *Hazz al-quḥūf fī ʔarḥ qaʔid Abī ʔādūf*; Vollers (1895) is a descriptive grammar of Egyptian Arabic at the close of the nineteenth century, and Elias & Elias (1981) an English–Egyptian Arabic vocabulary and phrasebook first released ca. 1899; Abdel-Massih et al. (2009) is a reference grammar of modern Egyptian Arabic.

Having established these understandings of grammaticalization and its component processes, we now turn to the proposal that specific grammaticalization pathways are able to be shared between interacting languages and dialects: the aforementioned CIG.

### 1.3 Contact-induced grammaticalization

The most fully elaborated theoretical model proposed for CIG is that of Heine & Kuteva (2003; 2005). This model represents the phenomenon by which “a grammaticalization process [...] is transferred from the model (M) to the replica language (R),” without corresponding transfer of any actual phonological form (2003: 539). As paraphrased and clarified by Law (2014: 215), this occurs when one language, “the ‘replica language,’ develops a feature observed in another language, the ‘model’ language, but goes through a path of universal development using resources internal to the replica language.” The result is such as that seen in the Basque innovation of an allative preposition from the noun *buru* ‘head’ and a perfect tense formed with the lexical verb *ukan* ‘have’, apparently influenced by parallel grammaticalizations in neighboring varieties of Romance (Haase 1992 *apud* Heine & Kuteva 2003: 550). Such an effect is proposed to be actuated according to the following model (Heine & Kuteva 2003: 539):

1. Speakers of language R notice that in language M there is a grammatical category  $M_x$ .
2. They develop an equivalent category,  $R_x$ , using material available in their own language (R).
3. To this end, they replicate a grammaticalization process they assume to have taken place in language M, using an analogical formula of the kind  $[M_y > M_x] = [R_y > R_x]$ .
4. They grammaticalize  $R_y$  to  $R_x$ .

This proposal for the diffusion of parallel grammaticalization trajectories between linguistic varieties is presaged by Bisang’s (1998) observation of the potential synergy between grammaticalization, which he considers to be a primarily construction-based process, and previously observed forms of contact-induced structural convergence. The phenomenon has also been influentially described by Dahl in the form of “gram families”, consisting of groups of “grams [grammaticalized items] with related functions and diachronic sources that show up

in genetically and/or geographically related groups of languages, in other words, what can be assumed to be the result of one process of diffusion” (Dahl 2001: 1469). Heine & Kuteva draw heavily on Dahl’s theorizations, though they diverge from him in a few critical ways. First, they are significantly more conservative than Dahl in identifying examples of the phenomenon, insisting on corroborating evidence of language contact in order to posit CIG rather than inductively inferring its occurrence given genetic relatedness or proximity. Second, they do not necessarily attempt to link multiple replications of the same grammaticalization pathway into “one process of diffusion,” but instead prefer to treat them as individual instances of contact between participating languages.

Further, Heine & Kuteva’s model is primarily situated in the context of contact between genetically distinct languages. Regarding the occurrence of CIG between related language varieties or dialects, Dahl sees such scenarios as generating the bulk of evidence for the phenomenon: “in the majority of all such cases [of areally diffused grams], the languages involved are more or less closely related” (2001: 1469). Heine & Kuteva are wary of such identifications. Critically, however, their reasons for being so are methodological rather than theoretical. In their analysis, they choose to rely on the principle of genetic patterning as “an empirically well-founded tool for identifying cases of contact-induced linguistic transfer” (2005: 33–34), meaning that examples of CIG between unrelated languages are often easiest to identify and defend and thus have been favored in the effort to present an unambiguous account. Regarding the broader occurrence of the phenomenon, however, they state that “genetic relationship is entirely irrelevant” (2005: 184) and that CIG may occur between related languages just as it does between unrelated ones. They remain, though, more careful than Dahl to set apart cases attributable to inheritance of any stage of the grammaticalization chain from a common ancestor, which could lead to a superficially similar result not in fact dependent on any degree of contact. Along the same lines, Law (2014) reminds us that when dealing with closely related languages the possibility of drift or typological poise precipitating parallel development rises dramatically in likelihood. Thus, the analyst must be stringent in linking proposed cases of CIG to cross-linguistically attested paths and parameters of grammaticalization and not to the local idiosyncrasies of a given language family or subgroup.

To Heine & Kuteva, CIG is unambiguously situated in terms of Van Coetsem’s (1988; 2000) dichotomy between source language (SL) and recipient language (RL) agentivity: the four-stage model of replication presented above clearly identifies speakers of the RL as the agents of contact-induced change in this instance. This judgment has opened the proposal to major critique, as several key theorists maintain that structural pattern replication of the kind required for CIG

is only possible in a scenario of SL agentivity. Ross (2007), for example, views the phenomenon as part of a broader process of bilingual calquing (cf. Manfredi, this volume), involving the subconscious imposition of the functional range of an SL item onto its RL equivalent, followed secondarily by the processes Heine & Kuteva attribute to grammaticalization but which Ross views as the natural result of increases in frequency and automatization stemming from the RL item's new functionality. Ross asserts that "one cannot reasonably argue" for Heine & Kuteva's construal of CIG as an RL-agentive direct replication of a grammaticalization process because of his conviction that the phenomenon is "largely driven by effort reducing practices of which speakers are only marginally aware" (2007: 135).

Matras (2009; 2011), however, supports Heine & Kuteva's initial characterization by arguing for RL-agentivity in his own recent accounts of CIG, which provide more attention to the role played by the individual bilingual in the phenomenon's actuation. He cites the individual's communicative imperatives and creative impulse as the primary force driving the replication of grammaticalization processes, as speakers actively borrow from constructions they control in one of their languages, as a source of expressive innovation in the other, limiting this transfer solely to "pattern" out of respect for the norms of the distinct speech communities in which they operate. Matras' account thus has the benefit of aligning with the motivating forces theorized to obtain for grammaticalization processes more generally. As described by Lehmann in his consideration of grammaticalization's communicative/pragmatic dimension:

To the degree that language activity is truly creative, it is no exaggeration to say that languages change because speakers want to change them ... they do not want to express themselves the same way they did yesterday, and in particular not the way that somebody else did yesterday (Lehmann 1985: 315).

Building upon this position, it holds that in scenarios of language or dialect contact innovating speakers may very well wish to express themselves the same way somebody else did yesterday if the means of expression involved are novel to a distinct speech community with which they are interacting today. This synergy with less controversial understandings of grammaticalization outside the context of language/dialect contact provides a viable counterpoint to the skepticism voiced by Ross, and strongly recommends the association of CIG with RL-agentivity.

In the case of contact between closely related varieties, this characterization of CIG may be further qualified. Under Matras' RL-agentive formulation, pattern replication occurs in the presence of constraints against the usage of an SL's forms due to speaker expectation and "language loyalty" among members of the RL community (2011: 283). In the broader language contact literature, such sociolinguistic constraints have been noted to play a role in pattern diffusion (e.g. Epps 2005), and presumably interact with speakers' judgments of interlocutors' perceived bilingual competency in favoring or disfavoring matter-based mixing or borrowing (cf. Grosjean 2001). In contexts where mutual comprehensibility is a less salient concern, the drivers of pattern vs. matter-based innovation would be expected to be almost purely sociolinguistic and pragmatic, and are perhaps most fruitfully understood through the lenses of indexicality (Silverstein 2003) and focus at the level of the speech community (Le Page & Tabouret-Keller 1985) rather than as a desire to adhere to a reified linguistic code. Such is the state of affairs most likely to obtain in the case of CIG between neighboring varieties of Arabic, to which we now turn in detail.

## **2 CIG in the development of Arabic future tense markers**

### **2.1 Methods of investigation**

In the following subsections, I present evidence for the role of CIG as the primary mechanism underlying the development and distribution of future tense markers in the modern Arabic dialects. The data considered is drawn from a survey of eighty-one geographic sample points spanning the contiguous Arabic-speaking world, based on a total of eighty-eight descriptive sources.<sup>2</sup> This sample was constructed as part of the broader investigation of CIG between Arabic dialects presented by Leddy-Cecere (2018), which investigates the role of CIG in the development of a number of morphosyntactic features in modern Arabic varieties, including future tense markers, genitive exponents, and temporal adverbs meaning 'now'. A discussion of data and findings for the first of these features is the focus of this chapter, and these shall be seen to argue strongly for the identification of CIG as a key force in shaping the evolution of modern Arabic dialects. Readers are encouraged to refer to Leddy-Cecere (2018) for additional examination and expansion of the points to follow.

---

<sup>2</sup>Details of sample composition and sources consulted, as well as the selection criteria for each, may be found in Leddy-Cecere (2018: 34–35, 43–46).



To begin, I first examine the complete set of specific grammaticalizations of future tense markers attested by the Arabic dialect data. These have been identified via the observation of concurrent processes of desemantization, extension, decategorialization and erosion (as described in §1.2). These individual instances are further sorted into higher level groupings by grammaticalization path: future tense markers deriving from motion verbs meaning ‘go’, for example, as against those deriving from purposive constructions, etc. Special attention is paid to those specific grammaticalization paths represented by multiple evolutions involving distinct etyma, thus identifiable as potential candidates for the products of replication through CIG. As a final step in the evaluation, the geographic incidence of forms representing such multiply attested paths is considered to assess whether their modern distribution is consistent with a historical account of diffusion via contact. This latter portion of the analysis will be presented in §2.3, following the complete accounting of grammaticalized forms provided in §2.2 immediately below.<sup>3</sup>

## 2.2 Grammaticalizations of Arabic future tense markers by grammaticalization path

### 2.2.1 Futures from ‘go’ (FUT < GO)

Grammaticalizations of future tense markers from forms of lexical verbs meaning ‘go’ are well represented in the Arabic data. This grammaticalization path is widely attested cross-linguistically, providing one of the major sources for the development of future tense markers worldwide (Bybee et al. 1994; Heine & Kuteva 2002). Grammaticalizations of specific items observed in the cross-dialectal Arabic sample are described below.

\*rāyih:

Future markers representing grammaticalized forms of an active participle \*rāyih ‘going’ are found across a broad east-west swath of the Arabic-speaking world, extending from southern Iraq in the east to Algerian territory in the west. Differing degrees of grammaticalization are attested, with some forms maintaining full phonological integrity and categorial membership (e.g. Basra *rāyih*; Mahdi 1985, which displays adjectival gender/number agreement with its subject) and others showing dramatic erosion and of loss of morphosyntactic autonomy (including Cairo *ha- ~ ha-*,

---

<sup>3</sup>For further discussion and justification of each stage of this heuristic, see Leddy-Cecere (2018: 36–43, 209–214).

as described in §1.2). Semantically, some forms, such as Algiers *rāḥ* and Jerusalem *rāyih* ~ *rāḥ* ~ *ḥā-*, are recorded as expressing a value of immediate future or future intent (Boucherit 2011; Rosenhouse 2011), while the majority are associated with a meaning of general futurity.

\**yādī*:

Grammaticalized forms of the active participle \**yādī* ‘going’ are common future markers throughout much of Morocco and adjacent regions. Reduced and invariant forms often co-exist alongside less grammaticalized reflexes, thereby attesting discrete links in an increasingly advanced grammaticalization chain, as seen in Casablanca *yādī* ~ *ya-* (Caubet 2011).

\**māšī*:

Grammaticalized future markers deriving from the active participle \**māšī* ‘going’ occur in two distinct geographic pockets, one centered in north-central Morocco and the other in Tunisia and eastern Algeria. In addition to more predictable effects of phonetic erosion and decategorialization, several forms from the latter area display a further example of irregular sound change with the sporadic denasalization of \*/m/ > /b/, as in Sousse *māš* ~ *bāš* (Talmoudi 1980).

\**sāyir*:

Alone in the sample, Maltese attests a future marker deriving from a grammaticalized form of the active participle \**sāyir* ‘going’. This can be found in both an inflecting form *sejjer* (SG.M)/*sejra* (SG.F)/*sejrin* (PL) and as the more grammaticalized, invariant forms *ser* and *se* (Vanhove 1993).

### 2.2.2 Futures from ‘want’ (FUT < WANT)

Grammaticalizations from source constructions indicating desire or volition are another cross-linguistically common origin for future tense markers (Bybee et al. 1994; Heine & Kuteva 2002), and are similarly widespread to their FUT < GO counterparts in the Arabic dialect data. Specific grammaticalizations are discussed below.

\**yabyā* ~ *yabyī*:

Grammaticalized forms of the imperfect verb \**yabyā* ~ *yabyī* ‘want’ serve as future markers across a large portion of the Arabic-speaking world, stretching from the Arabian Peninsula across the Red Sea to the greater Sudanic area, and then northward through present-day Libya. While many Arabic varieties attest only a highly grammaticalized, reduced form of the

item (e.g. Abu Dhabi *b-*; Qafisheh 1977), other dialects display direct evidence of multiple stages of the grammaticalization chain, e.g. Ḥarb *yabyā* ~ *yabā* ~ *ba-* (Il-Hazmy 1975).

\**biddu* ~ *widdu*:

Future tense markers arising from grammaticalizations of \**biddu* ~ *widdu* ‘want’ are found throughout the broader Levantine area. In their most phonetically reduced forms (e.g. Soukhne *b-*; Behnstedt 1994), they are often superficially indistinguishable from the highly grammaticalized products of \**yabyā* ~ *yabyī* discussed above; several dialects, however, provide clear evidence for a distinct chain of development, such as Jebel Ansariye *baddo* ~ *bado* ~ *b-* (Lewin 1969) and Cilicia *baddu* ~ *baddi-* ~ *bad-* (Procházka 2011). In some varieties, grammaticalizations of \**biddu* ~ *widdu* operate alongside other markers of future tense to designate a more specified value: Damascus *bəddo* ~ *b-* is often reported to denote a modal value of possible or planned future, as opposed to the \**rāyiḥ*-derived forms *raḥ* ~ *ḥa-* which indicate a higher degree of certainty or expectation (cf. Lentin 2011). In other dialects, these forms would appear to have further desemantized and extended to a value of more general futurity. Future investigation is needed into the degree to which reduced reflexes of \**biddu* may have merged in mental representation with the continuous aspect marker *b-* present in many of the same varieties; relevant parallels might be drawn with scenarios of near homophony like that found in the dialect of Dhofar, in which continuous *bi-* exists alongside future *bā-* (< \**yabyā* ~ *yabyī*; Davey 2016).

\**yišā*:

In a number of Yemeni dialects, the future tense marker may be traced to a grammaticalized form of \**yišā* ‘want’. It is notable that in cases such as Sana’a *ša-* this form is used only with the first person singular verb (Watson 1993); in such circumstances, it is possible that its ultimate source should be more properly identified with \**ašā* ‘I want’.

\**ydawr*:

Varieties belonging to the Ḥassāniyya dialect complex of Mauritania and neighboring Mali are recorded as utilizing a grammaticalized form of the verb \**ydawr* ‘want’ with a following imperfect verb to denote a value of intentional future. This grammaticalization is relatively light, consisting primarily of desemantization and extension with little in the way of de-categorialization or erosion: Nouakchott *yo:r*, for example, denotes future intent while continuing to operate morphosyntactically as a fully inflected finite verb (Taine-Cheikh 2011).

\*byā:

Dialects of southern Morocco and southwestern Algeria occasionally attest grammaticalized forms of \*byā ‘want’ expressing a future tense value. Though lexically similar in origin to the grammaticalizations based on \*yabyā ~ yibbā ~ yibbī discussed above, the phonological shape of these items (e.g. Marrakech *bya*: ~ *ba*-; Sánchez 2014) recommends an identification of their source in the perfect stem \*byā, which is the typical means for expressing ‘want’ in this area.

### 2.2.3 Futures from ‘come’ (FUT < COME)

Another cross-linguistically common path of future tense grammaticalization, that involving verbs meaning ‘come, return’ (Bybee et al. 1994; Heine & Kuteva 2002), is represented in the Arabic data by markers originating from a single source etymon, \*ʕād ‘return’.

\*ʕād: Future tense markers traditionally identified as grammaticalized forms of \*ʕād ‘return’ are attested in three locations in the cross-dialectal survey: Yemen, Upper Egypt and interior Tunisia. The forms found in Tunisia and Egypt, Tozeur *ʕa*- and Aswan *ʕa*- (Saada 1984; Schroepfer 2019), are highly reduced, and thus difficult to ascribe definitively to a specific source. It is notable that in both of these dialects the markers in question vary with a ‘go’-derived future *ha*- and could thus plausibly represent an erosion of the latter in the form of a sporadic lenition of /ħ/ > /ʕ/ (not to mention that Aswan *ʕa*- on its own might be linked to local *ʕāyiz* ~ *ʕāwiz* ‘want’). At least in the case of the Yemeni forms, however, an origin in \*ʕād seems clear, as reduced forms such as Sana’a *ʕā*- display an allomorph *ʕad*- in prevocalic contexts (Watson 1993).

### 2.2.4 Futures from purposive constructions (FUT < PURP)

A further source of future tense markers in the Arabic data involves the grammaticalization of purposive operators. This path is not widely discussed in the cross-linguistic grammaticalization literature, though intriguingly the reverse trajectory, that of PURP < FUT, is noted (Bybee et al. 1994). The primary difficulty would seem to rest in the identification of a clear process of desemanticization, as it is difficult to judge precisely which function between FUT and PURP is more concrete/abstract than the other. Despite this, the occurrence of extension, decategorialization and erosion in the Arabic forms seems to recommend their identification as products of a grammaticalization process.

\*ḥattā:

Grammaticalizations of \*ḥattā ‘in order to’ are used to indicate future tense in areas of northern Mesopotamia, the coastal Levant, and Oman. In terms of geographic distribution and the specific path of phonetic erosion followed, it may be possible to recognize Levantine and Mesopotamian forms like Cypriot Maronite *tta-* and Mosul *də-* (Jastrow 1979; Borg 1985) as representing a single historical innovation, though Oman *ḥa-* ~ *ha-* is more likely an independent development. In the Omani case, the attested use of *ḥa-* with purposive meaning recommends a source in \*ḥattā rather than GO-future \*rāyih: *šrab ḥa-turwe!* ‘Drink so your thirst be quenched!’ (Reinhardt 1894: 276).

### 2.2.5 Futures from ‘to busy oneself with’ (FUT < VERB OF ACTIVITY/PREPARATION)

A small number of Arabic dialects utilize a future tense marker seeming to derive from a grammaticalized form of a verb meaning ‘to busy oneself’. Such a path of development is not discussed in the cross-linguistic literature on grammaticalization, but perhaps has a counterpart in the use of grammaticalized Southern American English *fixing to* ~ *fixin’ a* ~ *fi’na* to express proximate futurity (cf. Wolfram & Schilling-Estes 1998). In any case, obvious desemanticization, extension, decategorialization and erosion of the source form indicate a clear example of grammaticalization here.

\*lāhi:

In the Ḥassāniyya dialects of Mauritania, Mali and far southern Morocco, the future tense marker derives from a grammaticalized form of \*lāhi, itself the active participle form of the verb *lha* ‘to busy oneself’. Decategorialization is attested in all cases by the lack of adjectival agreement marking predicted for the original participial, and in at least some varieties phonetic erosion is evidenced as well: Mali *lāhi* ~ *lā* (Heath 2003).

## 2.3 Evidence of replication and diffusion via contact

Of the five grammaticalization paths for Arabic future markers presented above, two merit closer examination in the search for evidence of CIG: those of FUT < GO and FUT < WANT. These paths are identified due to the fact that each is represented in the data by multiple, parallel realizations arising from etymologically distinct but semantically and functionally analogous sources. Such a state of

affairs plausibly reflects the result of continued processes of replication, whereby a grammaticalization process occurring in one Arabic variety is transferred to another and recreated using native etymological material.

Both paths identified, however – together representing the great majority of future tense markers attested in the sample – are also extremely common cross-linguistically, and could, in principle, have fed multiple independent developments instantiated across the modern Arabic dialect continuum. Key to selecting between an analysis of CIG and one of repeated, internally-motivated grammaticalization is the factor of geography, as in the absence of fine-grained historical sociolinguistic data (see §1.3) this is perhaps the most reliable proxy in positing the feasibility of historical contact between dialects. In the case of CIG, analogous grammaticalization processes ought to be positioned in a geographically contiguous (or near contiguous) bloc, consistent with a history of diffusion via contact between speakers of neighboring dialects. In a scenario of independent development, on the other hand, one should expect the various grammaticalizations to be more or less randomly distributed across the map, equally likely to occur in any individual dialect considered.

The geographic incidences of the members of the FUT < GO and FUT < WANT paths both clearly align with the contiguous profile anticipated for the results of CIG. All realizations of FUT < GO future tense markers described connect geographically with other members of the bloc. The large eastern and central zone of \*rāyih futures, encompassing southern Mesopotamia, much of the Levant, and the Nile Valley, stretches westward across Libya (where \*rāyih-derived forms are recorded alongside \*yibbi-based WANT-futures) to include most of Algeria. Directly adjacent to this North African arm of the \*rāyih forms are found grammaticalizations of \*yādī in Morocco and of \*māši in Tunisia. Further neighboring or co-territorial with the latter two areas are a second set of \*māši-based forms in northern Morocco and the Maltese \*sāyir-derived future tense marker, thus completing the connected geographic trend. Future markers representing the path FUT < WANT display a similar spatial contiguity. Grammaticalizations of \*biddu ~ widdu in the Levant stretch to reach those of \*yabyā ~ yabyī present in the Arabian Peninsula. These in turn span the Red Sea across to the greater Sudanic area and northward through the central Sahara into Libya. Moving to the west and southwest of this zone, the next future markers encountered include grammaticalizations of \*byā and \*ydawr, respectively. Rounding out the set, forms derived from \*yišā exist in close proximity to analogous \*yabyā ~ yabyī futures in Yemeni territory. While the integrity of this WANT-future bloc may seem to be challenged by natural features such as the Red Sea and the Sahara Desert, historical and anthropological investigations of the regions in question have rather

shown persistent social and cultural connectivity across these would-be barriers (Lydon 2009; Power 2012). This evaluation is supported by the distribution of additional Arabic dialectological isoglosses extending beyond the discussion of CIG.

The geographic contiguity displayed by the representatives of both the FUT < GO and FUT < WANT pathways favors an interpretation of areal diffusion over one of independent, internally motivated occurrence (of the type perhaps evidenced by the more scattered distributions of the sole representatives of FUT < COME and FUT < PURP). The optimal account for the development of the modern Arabic GO- and WANT-futures, together representing the greater part of future tense markers attested in the data, is thus one by which grammaticalization processes leading to the development of new future tense markers have repeatedly been subject to transfer and replication between speakers of neighboring dialects. A CIG-driven analysis such as this has the benefit of accounting for both the development of individual dialect forms and more global trends in source semantics and geographic incidence, and offers a theoretically unified interpretation of the Arabic data obtaining on multiple scales.

### 3 Conclusion

The analysis summarized above has demonstrated the significant explanatory power of CIG as an account for the development of Arabic future tense markers. Additional proposed occurrences of CIG between Arabic dialects, pertaining to genitive exponents and temporal adverbs meaning ‘now’, are identified and examined in Leddy-Cecere (2018). Together with the future tense data discussed here, these call for corroboration and refinement at the hands of future investigators.

Should further examination provide evidence for a widespread history of CIG between Arabic dialects, this finding could prove instrumental in satisfactorily accounting for a number of so-called “pluriform” developments which have repeatedly vexed students of Arabic dialectology. Defined by Versteegh as functionally analogous but etymologically disparate developments for which “a general trend ... has occurred in all Arabic dialects, and an individual instantiation of this trend in each area,” dialect contact has most often been dismissed as a causal mechanism for these innovations due to a belief that “typically dialect contact leads to the borrowing of another dialect’s markers, not to the borrowing of a structure, which is then filled independently” (2014: 146). CIG provides a theoretical mechanism by which precisely such borrowing and filling may occur, and as such offers the dialectologist a novel analytical tool in the elucidation of structural transfer and diffusion between Arabic varieties.

A critical open question in the application of CIG to the Arabic context, as well as in study of the phenomenon more generally, lies in the problem of agentivity and actuation (as discussed in §1.3). Here, too, further accrual of Arabic data has the potential to inform broader domains of inquiry. If Arabic is established as a productive ground for the study of CIG and significant cases of the historical transfer of grammaticalization pathways between dialects are brought to light, it stands to reason that the same societal and linguistic forces which have motivated these to take place may still be in force, and that observation of synchronic Arabic dialect interaction represents a singular opportunity to catch newly occurring instances of CIG “in the act” and to observe their progress in real time (for at least one such attempt already presented, see Abuamsha 2016). Studies of this type will enable linguists to add critically lacking synchronic data to their sociolinguistic and psycholinguistic analyses of CIG, and so elaborate and strengthen ongoing theorizations of a revelatory new dimension of contact-induced language change.

## **Further reading**

For a complete theoretical discussion of contact-induced grammaticalization, see Heine & Kuteva (2003; 2005). The former work is an article-length sketch of the proposal and is valuable as a direct and concise reference, while the latter provides a more elaborated description with additional linguistic examples. Matras (2009) provides valuable commentary and critique of Heine & Kuteva’s work while simultaneously extending exploration to the psycholinguistic and sociolinguistic dimensions of CIG.

For an overview of grammaticalization processes in the development of Arabic future tense markers (though without reference to contact) see Stewart (1998). A more detailed treatment of CIG processes in the development of the Arabic future markers and other morphosyntactic features may be found in Leddy-Cecere (2018).

## **Acknowledgements**

I am grateful to Drs. Kristen Brustad and Daniel Law for their contribution to the broader work which this chapter reflects, as well to the participants and organizers of the workshop “Arabic and Contact-Induced Change” at the 23rd International Conference on Historical Linguistics for their insightful feedback relating to the ideas discussed here.



This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1110007. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

## Abbreviations

CIG	contact-induced grammaticalization	PL	plural
F	feminine	PURP	purposive
FUT	future	RL	recipient language
invar.	invariant	SG	singular
M	masculine	SL	source language

## References

- Abdel-Massih, Ernest, Zaki Abdel-Malek, El-Said Badawi & Ernest McCarus. 2009. *A reference grammar of Egyptian Arabic*. Washington, DC: Georgetown University Press.
- Abuamsha, Duaa. 2016. The future marker in Palestinian Arabic: An internal or contact-induced change? In Lindsay Hracz (ed.), *Proceedings of the 2016 Annual Conference of the Canadian Linguistic Association*.
- Andersen, Henning. 2008. Grammaticalization in a speaker-oriented theory of change. In Thórhallur Eythórsson (ed.), *Grammatical change and linguistic theory: The Rosenthal Papers*, 11–44. Amsterdam: John Benjamins.
- Behnstedt, Peter. 1994. *Der arabische Dialekt von Soukhne (Syrien). Teil 2: Phonologie, Morphologie, Syntax. Teil 3: Glossar*. Wiesbaden: Harrassowitz.
- Behnstedt, Peter & Manfred Woidich. 2005. *Arabische Dialektgeographie: Eine Einführung*. Leiden: Brill.
- Bisang, Walter. 1998. Grammaticalization and language contact, constructions and positions. In Ana Giacalone-Ramat & Paul Hopper (eds.), *The limits of grammaticalization*, 13–58. Amsterdam, Philadelphia: John Benjamins.
- Borg, Alexander. 1985. *Cypriot Arabic: A historical and comparative investigation into the phonology and morphology of the Arabic vernacular spoken by the Maronites of Kormakiti village in the Kyrenia district of north-west Cyprus*. Stuttgart: Deutsche Morgenländische Gesellschaft.
- Boucherit, Aziza. 2011. Algiers Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.

- Bybee, Joan. 2003. Cognitive processes in grammaticalization. In Michael Tomasello (ed.), *The new psychology of language*, vol. 2, 145–167. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago, London: Chicago University Press.
- Caubet, Dominique. 2011. Moroccan Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.
- Claudi, Ulrike & Bernd Heine. 1986. On the metaphorical basis of grammar. *Studies in Language* 10(2). 297–335.
- Dahl, Östen. 2001. Principles of areal typology. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, 1456–1470. Berlin: Walter de Gruyter.
- Davey, Richard J. 2016. *Coastal Dhofari Arabic: A sketch grammar*. Leiden: Brill.
- Davies, Humphrey. 1981. *17th-century Egyptian Arabic: A profile of the colloquial material in Yūsuf al-Širbīnī's Hazz al-quhūf fī šarḥ qaṣīd Abi Šādūf*. Berkeley: University of California. (Doctoral dissertation).
- Elias, Elias & Edward Elias. 1981. *Egyptian Arabic manual for self study*. 19th edn. Cairo: Elias Modern Publishing House.
- Epps, Patience. 2005. Areal diffusion and the development of evidentiality: Evidence from Hup. *Studies in Language* 29(3). 617–650.
- Grosjean, François. 2001. The bilingual's language modes. In Janet Nicol (ed.), *One mind, two languages: Bilingual language processing*, 1–22. Malden: Blackwell.
- Haase, Martin. 1992. *Sprachkontakt und Sprachwandel im Baskenland: Die Einflüsse des Gaskognischen und Französischen auf das Baskische*. Hamburg: Buske.
- Hansen, Maj-Britt Mosegaard & Jacqueline Visconti. 2009. On the diachrony of “reinforced” negation in French and Italian. In Corinne Rossari, Claudia Ricci & Adriana Spiridon (eds.), *Grammaticalization and pragmatics: Facts, approaches, theoretical issues*, 137–171. Bingley: Emerald Group.
- Haspelmath, Martin. 1998. Does grammaticalization need reanalysis? *Studies in Language* 22(2). 271–287.
- Heath, Jeffrey. 2003. *Hassaniya Arabic (Mali): Poetic and ethnographic texts*. Wiesbaden: Harrassowitz.
- Heine, Bernd. 2004. Grammaticalization. In Brian Joseph & Richard Janda (eds.), *The handbook of historical linguistics*, 575–601. Oxford: Wiley-Blackwell.
- Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

- Heine, Bernd & Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3). 529–572.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Hopper, Paul & Elizabeth Traugott. 2003. *Grammaticalization*. 2nd edn. Cambridge: Cambridge University Press.
- Il-Hazmy, Alayan Mohammed. 1975. *A critical and comparative study of the spoken dialect of the Ḥarb tribe in Saudi Arabia*. Leeds: University of Leeds. (Doctoral dissertation).
- Jastrow, Otto. 1979. Zur arabischen Mundart von Mossul. *Zeitschrift für Arabische Linguistik* 2. 76–99.
- Law, Daniel. 2014. *Language contact, inherited similarity and social difference: The story of linguistic interaction in the Maya Lowlands*. Amsterdam, Philadelphia: John Benjamins.
- Le Page, R. B. & Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.
- Leddy-Cecere, Thomas. 2018. *Contact-induced grammaticalization as an impetus for Arabic dialect development*. Austin: University of Texas. (Doctoral dissertation).
- Lehmann, Christian. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e Stile* 20. 303–318.
- Lehmann, Christian. 2015. *Thoughts on grammaticalization*. Berlin: Language Science Press. DOI:10.17169/langsci.b88.98
- Lentin, Jérôme. 2011. Damascus Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.
- Lewin, Bernhard. 1969. *Notes on Cabali: The Arabic dialect spoken by the Alawis of “Jebel Ansariye”*. Stockholm: Almqvist & Wiksell.
- Lydon, Ghislaine. 2009. *On trans-Saharan trails: Islamic law, trade networks, and cross-cultural exchange in nineteenth-century western Africa*. Cambridge: Cambridge University Press.
- Mahdi, Qasim R. 1985. *The spoken Arabic of Baṣra, Iraq: A descriptive study of phonology, morphology and syntax*. Exeter: University of Exeter. (Doctoral dissertation).
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Matras, Yaron. 2011. Grammaticalisation and language contact. In Bernd Heine & Heiko Narrog (eds.), *Handbook of grammaticalisation*, 279–290. Oxford: Oxford University Press.

- Miller, Catherine, Enam Al-Wer, Dominique Caubet & Janet C. E. Watson (eds.). 2007. *Arabic in the city: Issues in dialect contact and language variation*. London: Routledge.
- Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.
- Power, Timothy. 2012. *The Red Sea from Byzantium to the Caliphate, AD 500–1000*. Cairo: The American University in Cairo Press.
- Procházka, Stephan. 2011. Cilician Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.
- Qafisheh, Hamdi. 1977. *A short reference grammar of Gulf Arabic*. Tucson, Arizona: University of Arizona Press.
- Ramat, Paolo. 2011. Adverbial grammaticalization. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of grammaticalization*, 502–510. Oxford: Oxford University Press.
- Reinhardt, Carl. 1894. *Ein arabischer Dialekt gesprochen in 'Omān und Zanzibar*. Stuttgart & Berlin: W. Spemann.
- Rosenhouse, Judith. 2011. Jerusalem Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.
- Ross, Malcolm. 2007. Calquing and metatypy. *Journal of Language Contact* 1(1). 116–143.
- Rubin, Aaron D. 2005. *Studies in Semitic grammaticalization*. Winona Lake, IN: Eisenbrauns.
- Saada, Lucienne. 1984. *Eléments de description du parler arabe de Tozeur (Tunisie): Phonologie, morphologie, syntaxe*. Paris: Librairie Orientaliste Paul Geuthner.
- Sánchez, Pablo. 2014. *El árabe vernáculo de Marrakech: Análisis lingüístico de un corpus representativo*. Zaragoza: Prensas de la Universidad de Zaragoza.
- Schroepfer, Jason. 2019. Notes on Aswan Arabic. Unpublished manuscript, University of Texas.
- Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication* 23. 193–229.
- Stewart, Devin. 1998. Clitic reduction in the formation of modal prefixes in the post-classical Arabic dialects and Classical Arabic “sa-/sawfa”. *Arabica* 45(1). 104–128.
- Taine-Cheikh, Catherine. 2011. Ḥassānīyya Arabic. In Lutz Edzard & Rudolf de Jong (eds.), *Encyclopedia of Arabic language and linguistics*, online edn. Leiden: Brill.
- Talmoudi, Fathi. 1980. *The Arabic dialect of Sūsa (Tunisia)*. Göteborg: Acta Universitatis Gothoburgensis.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.

- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Vanhove, Martine. 1993. *La langue maltaise: Etudes syntaxiques d'un dialecte arabe «périphérique»*. Wiesbaden: Harrassowitz.
- Versteegh, Kees. 2014. *The Arabic language*. 2nd edn. Edinburgh: Edinburgh University Press.
- Vollers, Karl. 1895. *The modern Egyptian dialect of Arabic: A grammar with exercises, reading lessons and glossaries*. Translated by Francis Burkitt. Cambridge: Cambridge University Press.
- Watson, Janet C. E. 1993. *A syntax of Ṣanʿānī Arabic*. Wiesbaden: Harrassowitz.
- Wolfram, Walt & Natalie Schilling-Estes. 1998. *American English: Dialects and variation*. Malden: Blackwell.

