

estudios de dialectología  
norteafricana y andalusí  
8 (2004), pp. 239-251

### WHAT'S WHERE AND HOW'S WHAT? INTERROGATIVES IN ARABIC DIALECTS

KEES VERSTEEGH

In many languages, interrogative pronouns and adverbs form a paradigmatic set etymologically, even though synchronically this is not always visible. In Germanic languages, for instance, interrogatives are often built with a prefix *w-* or *wh-*, and in Greek most of them contain *p-*, in Latin *qu-*. To some extent, one could probably claim that for the speakers of these languages these prefixes mark interrogativity. The regularity is, however, incomplete. Thus, in English along with *who* and *what* there is *how*, and in Latin along with *quando* and *quid* we have *ubi* and *cur*. For the learner of the language this means that, in spite of a certain paradigmatic regularity, there is still a lot of learning to do.

In some cases of language change, paradigmatic regularity is restored or enhanced by periphrasis of the interrogatives. This periphrasis consists in combining a uniform question particle with a nominal questioned element, for instance, 'what place?', 'what person?', 'what thing?', 'what manner?'. These newly developed, bi-morphemic forms are semantically more transparent in the sense of Seuren and Wekker (1986), since they can be broken down into a part functioning as a question particle (Q) and a part functioning as questioned word, i.e., Q-place, Q-thing, etc. Such a tendency towards language change is particularly prominent in processes of pidginization/creolization (Holms 1988:65, 87, 213), e.g., in Tok Pisin *wanem* 'what?' < Q + name, in Haitian French Creole *ki-gen* 'how?' < Q + genre, etc. Apparently, in the communicative situations in which pidgins and creoles emerge, and in which second language learning takes place in an improvised way, rather than under strict monitoring, there is a bonus on the enhancement of semantic transparency.

It should be emphasized here that pidgin/creole languages are not the only ones to have transparent interrogatives. According to Cysouw (2004:3), completely transparent systems are rare overall (he cites as examples Surinamese Sranan and Kenyan Pidgin Swahili), whereas incompletely transparent systems are somewhat more common (e.g., Ewe, Nambikuara and Pirahã). His sampling is synchronic, however, and what we are concerned with here is cases in which an opaque system is (partly) replaced by a transparent one.

The greater semantic transparency is also evident in the development of separate question particles for simple yes/no questions, possibly because in the most rudimentary forms of communication there is no change in word order between statements and questions, the latter being marked only by intonation. Question particles

are then indispensable for unambiguous marking of interrogative sentences. An example is the question particle in Principe Creole Portuguese (Holm 1988:213):

*kwá ci mesé â*  
 thing you want Q  
 'what do you want?'

A third development is that of phrasal interrogatives of the type 'who is that?', 'what is that?', for instance Tok Pisin *husat* or Jamaican *huu(-dat)* 'who?', both < *who's that*. These, too, may be said to enhance the pragmatic transparency of the utterance, by highlighting the interrogative. The utterance is split into two parts: rather than asking 'Who did this?', one asks 'Who is it? (He) did this' or possibly 'Who is it that did this?' (cf. French interrogatives of the type *qu'est-ce que . . . ?*).

Not all pidgin and creole language exhibit this type of transparent interrogative, and they do not exhibit it in all interrogatives, either. Muysken and Veenstra (1995:124-128) distinguish within this group between transparent systems, in which the question particle and the questioned element constitute semantic units; mixed systems, in which the question particle combines with an interrogative from the lexifier language; opaque systems, in which interrogatives from the lexifier language replace the creolized interrogatives; and atrophied systems, in which the question particle is dropped and only the questioned element remains. An example of an almost entirely transparent system is Chinese Pidgin English, with forms such as *wat-ting* 'what?' *wat-time* 'when?' *wat-side* 'where?', *wat-fashion* 'how?' (Muysken and Veenstra 1995:125). Mixed forms occur in almost all creoles, e.g., Jamaican *wen taym* 'when?', combining the questioned word *taym* with the interrogative *wen* from the target language. Such mixed forms may even involve mixing with an element from the substrate language or an element that is borrowed from another language. Muysken and Smith (1990:892) cite as an example the mixed forms in Berbice Dutch, in which *wa-anga* > *wanga* 'where?' seems to be a combination of Dutch *wat* > *wa* and Kalabari (an Eastern Ijo language) *angaa* 'place'. From Anatolian Arabic forms combining the Arabic question particle with a Turkish questioned word are reported, e.g., *ašyöl* 'how?' < Arabic 'what?' + Turkish *yol* 'way' (see below). Opaque forms are found in most pidgin and creole languages, for instance Tok Pisin *we(-stap)* 'where?' < *where* + *stop*; in some creolized languages almost the entire system is opaque, for instance, in the Arabic creole Nubi, spoken in Kenya and Uganda, in which most interrogatives derive from mono-morphemic forms, and in the English creole Saramaccan in Dutch Surinam. Atrophied forms occur in many creoles, e.g., in Sranan *pe* 'where?' < *o-pe* < Q + *presi* < *place* (Bruyn 1991), and in the Portuguese-based creole based Fad'Ambon (Post 1995:199) where *xa* 'what?' has developed from Q + *xa* < Portuguese *cosa* '(some)thing'. As we shall see below, it is also possible for atrophied forms to consist of the question particle only, the questioned word having been dropped, e.g. Egyptian 'ē 'what?' < 'ayyu šay'in.

Muysken and Smith (1990:901) conclude that "a large number of creoles have developed semantically transparent systems", but they admit that a number of questions are still open. For one thing, even in those languages in which the majority of the interrogatives is semantically transparent at the initial stage, the system may be subject to change at a later time. Bruyn (1991) has studied the development

of bi-morphemic interrogatives in Sranan (and English-based creole in Dutch Surinam) from the 18th century to the present time. She concludes that the original system was more transparent than the present one, since almost all interrogatives used to contain the question particle *hu-* (from English *who* or *how*). But in modern Sranan at least some of the interrogatives can be used optionally without the question particle, for instance (*o-*)*fa* 'how?' (Q-*fasi* < *fashion*) and (*o-*)*pe* 'where' (Q-*presi* < *place*). This development runs counter to the tendency towards transparency, since the forms without the question particle no longer belong to a paradigmatic set characterized by *hu-* (in older Sranan) or *o-* (in modern Sranan). Bruyn explains this by pointing out that the optional use of the question particle seems to be allowed only in those cases in which the questioned element is not homophonous with any free noun. Thus, for instance, in the interrogative *o-ten* 'when?' (Q-*ten* < *time*) the question particle is obligatory since the questioned element also occurs as a free noun meaning 'time'. In this way the principle of transparency is preserved, albeit indirectly (Bruyn 1991:45), because the question particle may be left out only in those cases in which no ambiguity arises.

The most important questions would seem to be, however, why some creoles developed almost completely transparent systems, while in others the system remained opaque. A partial answer to this question is given by Clements and Mahboob (2000). They point out that in situations of language contact there is always a process of what Thomason and Kaufman (1988) call 'mutual linguistic accommodation'. Thomason and Kaufman (1988:157) claim that:

"Acquisition of TL [target language] structures is primarily determined, for any language learning situation, by access to the TL and by motivation to learn it, and only secondarily by markedness and by the typological fit between TL structures and structures in the learners' language(s)".

This is translated by Clements and Mahboob (2000:462) into a hierarchy of factors to explain the possible outcome of the process of accommodation by the new learners of the target language:

- a. Access to the TL
- b. Motivation on the part of the shifting speakers to learn the TL in a context of creolization
- c. The typological distance that separates the languages in a given contact situation
- d. The tendency to use universally unmarked features in the creolization process".

By ordering the factors in this way, they have found a way to explain developments in the learners' variety of the language. If these have ample access to the target language and/or if they are motivated to shift completely to the target language, they will tend to take over wholesale the forms they are exposed to. Clements and Mahboob's approach effectively neutralizes the old opposition between substratal influence and autonomous development. According to them, substratal influence is only one of the factors to determine the outcome of the acquisition process. In this model it only comes to play a role if the typological distance

between the target language and the learners' language(s) is small. If the distance is increased, the likelihood of what they call 'universal features', for instance, a tendency towards semantic transparency, become operative. They apply this model to the emergence of interrogatives in a large number of English-, French, Portuguese- and Spanish-based pidgins/creoles and conclude (2000:492) that "if the languages in contact do not share a certain structure, then the structure of the question will be accounted for by the universal tendencies such as the drive for semantic transparency".

Inversely, a construction shared by two languages in contact has a higher chance to be preserved even if it is highly marked in the target language. This principle applies, for instance, to the choice between *in situ* interrogatives and *wh*-movement. The *in situ* position of the interrogative in Egyptian Arabic is often cited as a case of substratal influence from Coptic. Just as often, the validity of this case of alleged substratal influence is rejected. Behnstedt (*forthcoming*), for instance, points out that similar constructions are not completely unknown in Classical Arabic, even though they are highly marked. In his view, this means that it is unnecessary to assume Coptic influence to explain the position of the interrogative in Egyptian Arabic. In the model adopted here, the presence of the construction in Classical Arabic does not necessarily mean that influence from the substratal language is ruled out completely. In fact, both factors may have reinforced each other, according to the principle that if the typological distance between the two languages is only slight, marked constructions from the target language may be supported by similar constructions in the substratal language. In other words, if the new learners (in this case, of Arabic) are confronted with a choice between two alternatives, they are more likely to choose the one they are familiar with, even if it is marked in the target language, and even if this runs counter to universal tendencies towards semantic transparency.

When there is no variant to be favored on the basis of familiarity from the substratal structure, universal tendencies start to play a role. This appears to be the case of the bi-morphemic interrogatives of the kind described here. Their presence cannot be explained by the presence of such forms in any of the substratal languages (Coptic, Berber, Syriac), which all have mono-morphemic interrogatives. In Standard Arabic, just as in most other languages, bi-morphemic interrogatives do occur, but as stylistic variants along with the mono-morphemic forms. In English, 'what time?', 'what person?', 'what place?', 'what way?' are used by native speakers, as a simple google query immediately demonstrates. Although we have no figures about the use of such expressions in Classical Arabic, it may be assumed that forms like 'ayyu makānin, 'ayyu šay'in, 'ayyu mawḍi'in occurred regularly as stylistic variants. The difference between these variants and those in modern Arabic dialects is that in the latter bi-morphemic forms are used exclusively, in at least some dialects, for at least some of the interrogatives. This means that an explanation in terms of universal tendencies towards semantic transparency is appropriate. It is not my intention here to show that Arabic dialects went through a similar development as pidgin and creole languages (for that issue see Versteegh 2003), but to point out the relationship between periphrastic interrogatives and the communicative context of language acquisition. Since the first varieties of modern Arabic emerged in a situation of second language acquisition (cf. Al Sharkawi 2005), such an explanation has to take into account the linguistic accommodation involved in

the acquisition process and since substratal influence is unlikely in this case, the greater semantic transparency of the bi-morphemic variants must have played a role in the process of change.

Yet, the development of the Arabic dialects did not proceed along the same lines everywhere. In fact, examples are found of all four types of development of interrogative systems mentioned by Muysken and Veenstra (1995). In many dialects, there are at least some transparent, bi-morphemic interrogatives, whereas in other dialects reflexes of the Classical Arabic forms prevail, such as *mā*, 'ayna, *matā*, etc. These are opaque in the sense that their form cannot be broken down in smaller meaningful units. Along with the transparent and opaque forms, mixed or blended forms are also found in Arabic dialects. In these forms, the transparent questioned word is combined with a reflex of the Classical Arabic interrogative, in forms like 'aynaḥall 'where?' < 'ayna + maḥall 'place' in Antiochia Arabic or \*ayna mōda > Diyarbakir əndaḥ 'where?' < 'ayna + mawḍi' 'place'. Mixed forms of a different nature are also found, when the question particle is combined with a Classical Arabic interrogative, e.g., in Syrian Arabic 'ēmta 'when?' < 'ayy matā. Interrogatives of this kind, are relatively easy to explain as intrusion of the Classical Arabic form in the already existing transparent form. This phenomenon is well-known from other cases of language contact in which a target language continues to play a role or comes to play a role at a later stage, after new varieties of the language have already developed, as in the Jamaican example *wen taym* 'when?', cited above. Atrophied forms of the type referred to above also occur in the Arabic dialects. In these forms the Q-element is dropped from the questioned words. This represents a further erosion of the transparent system. In Moroccan Arabic we find, for instance, *šnū* 'who?' or an even shorter form, š- < 'ayyu šay'in huwa, in which the 'ayyu element is not visible anymore; in Maltese, 'what?' is *xi* < *šay*'. Somewhat different are cases where the question particle is the only part remaining after the deletion of the questioned word, as in Egyptian 'ē 'what?'. Such forms, too, are relatively easy to explain as a general erosion of frequently used function words. We have seen above in the case of Sranan that such an erosion can only take place if the remaining questioned word (Moroccan š-, Maltese *xi*-) is not homophonous with a free noun.

In discussing the development in the Arabic dialects I shall limit myself to the interrogatives 'who?', 'what?', 'where?', 'when?' and 'how?'. The situation with 'how many?' is more complicated, since as Manfred Woidich (p.c.) points out, the forms that are reflexes of *kam* do not have the same function as those deriving from some transparent form like \*'ayyu šay'in qadr 'what size, measure?', which is reflected in forms like š-qadr, š-gad in Mesopotamian Arabic, 'addi 'ē in Egyptian Arabic, or *áška* in Cypriot Arabic. The distinction between the two is probably identical with that between mass/count interrogatives, which is found in other languages as well.

	who?	what?	where?	when?	how?
Egyptian Arabic (Cairo)	<i>mīn</i>	<i>'ē</i>	<i>fēn</i>	<i>imta</i>	<i>izzāy</i>
Moroccan Arabic (Skūra, Aguadé & Elyacoubi 1995)	<i>škūn, mən</i>	<i>āš, əš; šnu</i> (Jewish)	<i>fīn</i>	<i>īnta</i>	<i>kīf, kīfaš</i>
Gulf Arabic (Holes 1990)	<i>mīn (man), mīnħu</i>	<i>šīnħu, wēš</i>	<i>wēn</i>	<i>mita</i>	<i>šlōn, kēf</i> (čēf)
Iraqi Arabic (Baghdad Muslim, Erwin 1963)	<i>mīnu</i>	<i>šīnu</i>	<i>wēn</i>	<i>šwakit</i>	<i>šlōn</i>
Iraqi Arabic (Baghdad Jewish, Mansour, <i>forth-</i> <i>coming</i> )	<i>māni</i>	<i>āškun</i>	<i>wēn</i>	<i>ēmta</i>	<i>āšlon</i>
Lebanese Arabic (Bišmizzīn, Jiha 1964)	<i>mīn</i>	<i>'aš, 'ayš</i>	<i>wayn</i>	<i>'aymtīn, 'aymta</i>	<i>kīf</i>
Syrian Arabic (Damascus, Grotzfeld 1965)	<i>mīn</i>	<i>šū, ('ēš)</i>	<i>wēn</i>	<i>'ēmta</i>	<i>šlōn</i>
Ḥassāniyya (Taine- Cheikh, <i>forthcoming</i> )	<i>mən</i>	<i>š(ə)-</i>	<i>mnāyn, wāyn</i>	<i>āyntā</i>	<i>škīv</i>
Cypriot (Borg1985)	<i>men</i>	<i>ayš, aš</i>	<i>'ayn, 'an</i>	<i>miten</i>	<i>āššik</i>
Uzbekistan Arabic (Fischer 1961)	<i>mīn</i>	<i>ēš</i>	<i>ēšāb</i>	<i>matā, īmit</i>	<i>ištūr</i>
Maltese	<i>mīn</i>	<i>xi</i>	<i>fejn</i>	<i>meta</i>	<i>kif</i>
Chadian Arabic	<i>yātu</i>	<i>šīnu</i>	<i>wēn</i>	<i>(mata), mitēn</i>	<i>kikēf, kikeff</i>
Arbīl (Jastrow1990)	<i>mani</i>	<i>aš, ašni</i>	<i>hēn</i>	<i>hēmate</i>	<i>āštōf</i>
Daragözü (Jastrow 1973)	<i>māni</i>	<i>šṭaba</i>	<i>əmma</i>	<i>čičax</i>	<i>əšma</i>
Bornu Arabic (Owens 1993)	<i>mīné</i>	<i>šīnú, šunú</i>	<i>yéen, wéen (weenu, yoonú)</i>	<i>matá, waqtéeš</i>	<i>kéef, kekkéef, kefkéef</i>
Nubi (Wellens2003)	<i>mu`nu</i>	<i>su`nu</i>	<i>(f)/(u)` wen, we`nu</i>	<i>mi`ten</i>	<i>ke`fin</i>

Table 1: Examples of interrogatives in some Arabic dialects

## WHO?

The interrogative 'who?' seems to be the one most resistant to periphrasis, which seems to be a general characteristic in the languages of the world: according to Cysouw (2004:2) only 5% of the world's languages have an analyzable (transparent) interrogative to ask about a person. Singer (1958:93ff) states that interrogative 'who?' in the Arabic dialects goes back to four different forms: *man*, *mīn*, 'ayyu šay'in huwa (e.g., Daḡina waššu), and 'ayyu šay'in yakūnu (e.g., Moroccan škūn). *mīn* is regarded by him as an original doublet of *man* (Singer 1958:129-130, 137). The forms with 'ayy šay'in are phrasal interrogatives, just like the form *minu* < *man huwa* in Šukriyya Arabic (Reichmuth 1983:116-117).

The phrasal origin of such phrasal interrogatives is still detectable in their syntactic use. They occur sentence-initially, or with a preposed relative clause, as in Šukriyya Arabic

*al-ḥaddasak minū*  
REL-told-you who  
'who told you?'

*al-fī-l-bēt minū*  
REL-in-ART-house who  
'who is in the house?'

This use goes back to a highlighting of the interrogative phrase or the subject phrase, i.e., 'the one who told you, who is he?' or 'who is he, the one who told you?'

In Chadian Arabic a different form for 'who?' is used, *yātu* (possibly derived from *ayy + t + hu*, Roth-Laly 1979:170-171), but it displays the same syntactic ordering; when used after a relative clause, *yātu* exhibits gender agreement with the subject of the relative clause:

*al-bi-thaddis yāti*  
REL-DUR-she.tells who  
'who [fem.] tells you?'

## WHAT?

For 'what?' almost all dialects have a form containing a reflex of Classical Arabic *šay* 'thing'. This is not in line with Cysouw's findings (2004:2), since only 5% of the world's languages have an analyzable form of the interrogative asking about a thing. The only place where reflexes of Classical Arabic *mā* are found is a large region in Yemen (Singer 1958:161). The distribution of *mā* and 'ayyu šay' informs in the Yemeni dialect area is illustrated strikingly by Map 59 in the dialect atlas of Yemen (Behnstedt 1985). The western part of this area has the reflexes of Classical Arabic *mā* for 'what?', whereas the eastern part has reflexes of 'ayyu šay'in 'what thing?'. In between are smaller areas in which the interrogative goes back to \**ma huwa* > *maw*, *mō*, *mū* or to \**mā hiya* > *mī*, sometimes reinforced with *-ši*, e.g., Ristāq *mhūšši* (Singer 1958:173). They have become completely gramma-

ticalized since they are also used when they do not function as subject or predicate, e.g., *mū gulk līn* 'what did you say to me?'

In general, the '*ayyu šay*' forms occur in two types of combination, both with and without final *-in*. According to Singer (1958:209) the forms with *-in* are older, for instance Maltese *šin*, Iraqi *šinu*, Tunisian *āšnūa*, Moroccan *ašnhūwa*, Andalusian *ašan*, *aššan*. In some dialects, forms with initial *w-* exist; according to Singer (1958:211) these are not the result of a combination with the conjunction *wa-*, but they go back to an alternation *w-~'*, e.g., in Aden Arabic *weš*, *waš*, *wuš*.

In some dialects, the questioned word has become atrophied through the loss of the question particle. This may be a development in real time, comparable to the development of the interrogative in Sranan: in Syrian Arabic the form '*ēš*' may be the older one (Grotzfeld 1965:23), which has given way to *šū* < '*ēš* + *hu*'. The inverse development has taken place in Egyptian Arabic, where the questioned word has dropped off (cf. Singer 1958:223-225 about the gradual replacement of '*ēš*' with '*ē*'). De Jong (2000:174, 290, 371) points out that in a number of Sinai Bedouin dialects (e.g., Rmēliyy, Smē<sup>c</sup>niyy, Biyyādiyy) '*ēš*' tends to be used sentence initially, '*ēh*' sentence finally. If this is a modern development, '*ēh*' may be the result of Cairene influence, in which case the sentence final position may reflect the *in situ* position of the Cairene interrogative.

#### WHEN?

For 'when?', Map 290 of the Syrian dialect atlas (Behnstedt 1997), provides an example of alternative forms restricted to small pockets: the overwhelming majority of the dialects has opaque reflexes of *matā* 'when?', but in the Qamišli area in the north-east transparent forms like *š-wakit*, *šōkt* < \*'*ayyu šay*' in *waqt* 'what time?' are found (cf. *áyšwaxt*, *ášwaxt*, etc., in Anatolian dialects, cf. Jastrow, *forthcoming*). A large number of places on the map displays mixed forms going back to combinations of '*ayy* + *matā*' > *aymta*, *amāta*, *ēmta*, etc. In Anatolian Arabic, mixed forms combining the Q-particle with the Classical reflex are found for 'when?', e.g., Mardin *áymate* (cf. Jastrow, *forthcoming*). For the Egyptian dialect area, Map 185 (Behnstedt and Woidich 1985) shows that most varieties have a reflex of \**matā* 'where?', but the northern part of the area has an expanded form *imta*, *aymta*, *ēmta* presumably derived from '*ayy matā*', in other words, a form combining the Classical Arabic interrogative with the Q-particle '*ayy*' (the boundary between these two areas lies north of il-Minya, Map 186). In a few places, a transparent form is found in which some form of *waqt* 'time' is combined with the question word: *waktē*, *waxtē* (in the oasis of Kharga and in the eastern Delta); both belong to the 'pre-Hilali' stage of Arabicization. In the rest of Egypt there is a clear distinction between Upper Egypt, with forms going back to Classical Arabic *matā*, and the rest of Egypt, including the Delta and Middle Egypt, with mixed forms going back to '*ayy* + *matā*': *imta*, *aymta*, *ēmta* (cf. Behnstedt and Woidich 1985, Map 185).

In North Africa, the situation is extremely varied, but Marçais (1977:253-254) makes a general division into reflexes of *waqt* and those of *matā*; the latter is mentioned for Tunisia and Libya (*amta*, *mta*). Marçais adds that *yəmta* is some-times



heard in Morocco. *əmtāš* is found in south Tunisia (Gabes), *mtēn* in Fezzan, *weynta*, *wēnta* in Algeria (the high plain of the Oranais and the Algérois). The forms related to *waqt* occur in the larger parts of Morocco and Algeria, according to Marçais also in Bedouin dialects; nonetheless, these must stem from sedentary dialects since in most cases they have either *q* or *k* (*wagtāh* is mentioned for the Bedouin dialects of the western Algérois and the Oranais, and *āna-wagt* for the Marāzīg in South Tunisia). In the dialect of Chauen, Moscoso (2003:189) notes the use of *fīwāx* for 'when?'.

#### WHERE?

For 'where?' most dialects have combinations of Classical Arabic 'ayna with *wa-* or *fī*. Mixed forms are found, for instance, in the Syrian dialect area (Behnstedt 1997, Map 286) in the form 'aynaḥalli 'where?' < \*'ayna + maḥall, which is found near Iskenderun (cf. Arnold, *forthcoming*, who documents *aynaḥall* for the Christian and Alawi dialect in Antiochia Arabic). In Anatolian Arabic forms going back to \*ayna mōḍa' 'which place?' occur, e.g., in Diyarbakir *əndaḥ* ~ *ənda*, Daragözü *əmmaḥ* ~ *əmma*. In the North African Arabic dialects, all forms mentioned by Marçais (1977:248) for 'where?' derive from combinations with 'ayna: *fāyn*, *wēn*, *lāyn*, *mnīn*, etc.

#### HOW?

For 'how?' the variation is extreme. Apart from the forms quoted in the tablet here are many more variants; for Anatolian Arabic alone Vocke and Waldner (1982:24-26) list *ašwan*, *aššon*, *āštawf*, *əštōr*, and even combinations with Turkish words such as *ašyōl* from Turkish *yol* 'way'. In Mardin *ašwan* is used for 'how?', the reflex of Classical Arabic *kayfa*, *kif*, is only preserved in greetings (*kif inti*) and as a conjunction (for the gramaticalization of *kayfa* in the Arabic dialects see Taine-Cheikh 2004). Borg (1985:146-148) mentions many examples of bi-morphemic interrogatives in Cypriot Maronite Arabic, e.g., *āššik* < 'ayy šakl 'which form?', and in Jewish Baghdadi Arabic *ašlōn* < 'ayy šay' lawn 'which color?'. In the Egyptian dialect area a division exists between forms going back to *kayfa* and those going back to *izzayy* < 'ēš zayy, with roughly the same geographical division as the alternatives for 'when?'. Interestingly, Kharga again has *izzayy*, along with *kīh*, whereas Dakhla and Farafra have *kīyyif* and *kayf*, respectively (Woidich, p.c.). All North African forms for 'how?' seem to go back to *kayfa*, but often combined with *-āš* (Marçais 1977:269-270).

What, then, could explain the development of this confusing pattern of transparent and opaque forms in the modern Arabic dialects? In view of the bewildering variety of the forms for the interrogatives in mainstream Arabic dialects it would be hard to reconcile the development of interrogatives with a monogenetic view of the emergence of the Arabic dialects. At least some of the transparent forms must have developed independently. If it is true that such a transformation of opaque into transparent forms originates in communicatively handicapped contexts, it may be assumed that they are vestiges of the earliest migration wave of Arab invaders, which led to a relatively quick process of Arabization. During this first stage of

the conquest, so-called sedentary dialects emerged in the larger urban centers of the conquered areas. In North Africa, such dialects are called 'pre-Hilali dialects'. During this period, what had thus far been a stylistic alternative of the Classical Arabic interrogatives had a much higher chance of being selected because it fitted the communicative situation. In the Maghrebi dialects, for instance, most of the forms for 'when?' that derive from *waqt* contain *q*, so that they probably belong indeed to the pre-Hilali stage. In other areas, too, the first conquest may be called a pre-Hilali stage. In Egypt, for instance, the distribution of the forms for 'when?' suggests that both transparent *waktē*, *waxtē* and the *izzayy* forms for 'how?' belong to this 'pre-Hilali' stage of Arabicization.

Later stages of Arabicization went at a much slower pace. The Hilali invasion of North Africa took centuries to run its course. In some of the areas invaded by Bedouin during this period, no second language acquisition was involved at all, since the Bedouin settled there and continued to speak their own dialect. If the original population eventually became Arabic-speaking, they adopted the language in the course of a slow process of acquisition, in which the speakers of the target language were far less a minority than had been the case during the early stages of Arabicization when Arabic-speakers in North Africa, for instance, may have been outnumbered at a ration of 1:100. The later Arabicization of the rural areas during the Hilali invasion involved a much higher degree of exposure to the target language and this, in turn, may have led to much better opportunity to accommodate to Arabic. In a similar way, Clements and Mahboob (2002:464-465) explain the difference between Cape Verdian Portuguese Creole and other creolized varieties of Portuguese by the large Portuguese presence in the Cape Verdian islands. In such a situation of maximum exposure and concomitant accommodation the continued use of directly borrowed interrogative forms from Classical Arabic is to be expected.

In some cases, even the transparent forms that had originally been introduced may have given way to opaque forms under the influence of Classical Arabic as used in the urban centers of administration and culture. If this happened, (some of) the original transparent forms were replaced by or combined with reflexes of the Classical Arabic forms. Examples of this process have been given above.

The acquisition process did not always lead to the emergence of bi-morphemic forms. Ki-Nubi did not develop transparent words, even though the communicative context in which the language originated would seem to have necessitated a high degree of transparency in the verbal communication. Nonetheless, there are other examples of Creole languages without or almost without bi-morphemic question words, e.g., in Fanakalo, a pidgin language based on the Bantu languages Zulu and Xhosa, and used in South-Africa in contacts between (black) workers and (white) management in the mines. According to Mesthrie (2002), the differences between Fanakalo and other pidgin languages are to be explained by the fact that it is a pidginized variety of a non-European language, which served as the target language for (some) Europeans in their contacts with the black population. Some of these circumstances also apply to Ki-Nubi: it, too, is based on a non-European language and its pidgin ancestor seems to have been used in the contacts between European officers of the Anglo-Egyptian army and Nubian recruits.

The main conclusion of the discussion above appears to be that the forms we have called transparent arose during the earlier wave, whereas the reflexes of Classical forms belong the second. Most Bedouin dialects do not have the extreme transparent forms, although they almost always have the transparent form for 'what?'. There are several avenues of further research that suggest themselves. In the first place, the structure of interrogation in the most primitive forms of foreigner talk remains unclear. As Al-Sharkawi (2005) has shown, the register to which the new learners were exposed was a systematically simplified form of Arabic. Questions figure prominently in speech addressed to foreigners (Long 1981). Yet, we do not know how users of this register re-structure the form of interrogatives and what information there is does not feature transparent forms of interrogatives. A more promising approach seems to be to focus on the presence in most languages of stylistic alternatives for the opaque forms and find out to what extent these were current in Classical Arabic.

Complete maps documenting the development of all interrogatives in all dialects are needed, since they might show patterns of distribution that may help to establish a correlation between the form of the interrogative and the settlement and Arabicization pattern. Clearly, without the tools of the dialects atlases it would be almost impossible to even begin to study this phenomenon. The lists of forms given by Marçais are difficult to interpret precisely because he does not provide precise maps on which the areas are shown. In this respect, Egypt, Yemen and Syria are areas where the painstaking work of Behnstedt (and for Egypt, Woidich) has made it possible to visualize the proliferation of the various forms on the maps.

#### LIST OF REFERENCES

- Aguadé, Jordi and Mohammad Elyaacoubi. 1995. *El dialecto árabe de Skûra (Marruecos)*. Madrid: Consejo Superior de Investigaciones Científicas.
- Al-Sharkawi, Mohammed. 2005. *Arabicization: A case of second language acquisition*. Ph.D. diss., University of Nijmegen.
- Arends, Jacques, Pieter Muysken and Norval Smith. 1995. *Pidgins and creoles: An introduction*. Amsterdam and Philadelphia: J. Benjamins.
- Arnold, Werner. *Forthcoming*. "Antiochia Arabic". *Encyclopedia of Arabic Language and Linguistics* I, 000. Leiden: E.J. Brill.
- Behnstedt, Peter. 1985. *Die nord-jemenitischen Dialekte*. I. Atlas. Wiesbaden: L. Reichert.
- .-. 1997. *Sprachatlas von Syrien. Kartenband + Beiheft*. Wiesbaden: O. Harrassowitz.
- .-. *Forthcoming*. "Coptic loanwords". *Encyclopedia of Arabic Language and Linguistics*, I, 000. Leiden: E.J. Brill.
- .-, and Manfred Woidich. 1985. *Die ägyptisch-arabischen Dialekte*. I. *Einleitung und Anmerkungen zu den Karten*. II. *Dialektatlas von Ägypten*. Wiesbaden: L. Reichert.
- 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-western Cyprus*. Stuttgart: F. Steiner.

- Bruyn, Adrienne. 1991. "Question words in 18th-century and 20th-century Sranan". *Historical Linguistics 1991. Papers from the 10th International Conference on Historical Linguistics, Amsterdam, 12-16 August 1991*, ed. by Jaap van Marle, 31-47. Amsterdam and Philadelphia: J. Benjamins.
- Clements, J. Clancy and Ahmar Mahboob. 2000. "Wh-words and question formation in pidgin/creole languages". *Language change and language contact in pidgins and creoles*, ed. by John McWhorter, 459-497. Amsterdam and Philadelphia: J. Benjamins.
- Cysouw, Michael. 2004. "Minimal systems of interrogative words". [email.eva.mpg.de/~cysouw/pdf/cysouwHANDMINQ.pdf](http://email.eva.mpg.de/~cysouw/pdf/cysouwHANDMINQ.pdf) [Downloaded 26.03.05.]
- Erwin, Wallace M. 1963. *A short reference grammar of Iraqi Arabic*. Washington, D.C.: Georgetown University Press.
- Fischer, Wolfdieter. 1961. "Die Sprache der arabischen Sprachinsel in Uzbekistan". *Der Islam* 36.232-263.
- Heine, Bernd and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Grotzfeld, Heinz. 1965. *Syrisch-arabische Grammatik*. Wiesbaden: O. Harrassowitz.
- Holm, John. 1988. *Pidgins and creoles. I. Theory and structure*. Cambridge: Cambridge University Press.
- Jastrow, Otto. 1973. *Daragözü, eine arabische Mundart der Kozluk-Sason- Gruppe (Südostanatolien): Grammatik und Texte*. Neurenberg: H. Carl.
- . 1990. *Der arabische Dialekt der Juden von 'Aqra und Arbîl*. Wiesbaden: O. Harrassowitz.
- . *Forthcoming*. "Anatolian Arabic". *Encyclopedia of Arabic Language and Linguistics* I, 000. Leiden: E.J. Brill.
- Holes, Clive. 1990. *Gulf Arabic*. London and New York: Routledge.
- Jiha, Michel. 1964. *Der arabische Dialekt von Bišmizzîn: Volkstümliche Texte aus einem libanesischen Dorf mit Grundzügen der Laut- und Formenlehre*. Beirut and Wiesbaden: F. Steiner.
- Jong, Rudolf de. 2000. *A grammar of the Bedouin dialects of the northern Sinai littoral: Bridging the linguistic gap between the Eastern and Western Arab world*. Leiden: E.J. Brill.
- Long, Michael H. 1981. "Questions in foreigner talk discourse". *Language Learning* 31.135-157.
- Mansour, Jacob. *Forthcoming*. "Baghdadi Arabic (Jewish)". *Encyclopedia of Arabic Language and Linguistics*, I, 000. Leiden: E.J. Brill.
- Marçais, Philippe. 1977. *Esquisse grammaticale de l'arabe maghrébin*. Paris: Maisonneuve.
- Moscoso García, Francisco. 2003. *El dialecto árabe de Chauen (N. de Marruecos)*. Cadiz: Universidad de Cádiz.
- Mesthrie, Rajend. 2002. "Fanakalo as an upside-down pidgin (abstract)". [http://www.nyu.edu/gsas/dept/lingu/events/urban\\_socioling/wgus02s/abstract/mesthrie.html](http://www.nyu.edu/gsas/dept/lingu/events/urban_socioling/wgus02s/abstract/mesthrie.html). [Downloaded, 26.03.05.]
- Muysken, Pieter and Norval Smith. 1990. "Question words in pidgin and creole languages". *Linguistics* 28.883-903.
- . and Tonjes Veenstra. 1995. "Universalist approaches". *Arends a.o.* (1995:121-134).

- Owens, Jonathan. 1993. *A grammar of Nigerian Arabic*. Wiesbaden: O. Harrassowitz.
- Post, Marijke. 1995. "Fa d'Ambu". Arends, Muysken and Smith (1995:191-204).
- Reichmuth, Stefan. 1983. *Der arabische Dialekt der Šukriyya im Ostsudan*. Hildesheim: G. Olms.
- Roth-Laly, Arlette. 1979. *Esquisse grammaticale du parler arabe d'Abbéché (Tchad)*. Paris: Geuthner.
- Seuren, Pieter and Herman Wekker. 1986. "Semantic transparency as a factor in creole genesis". *Substrata versus universals in creole genesis*, ed. by Pieter Muysken and Norval Smith, 57-70. Amsterdam and Philadelphia: J. Benjamins.
- Singer, H. 1958. *Neuarabische Fragewörter: Ein Beitrag zur historischen und vergleichenden Grammatik der arabischen Dialekte*. Diss. Universität Erlangen.
- Taine-Cheikh, Catherine. 2004. "De la grammaticalisation de 'comme' (comparatif) en arabe". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. by Martine Haak, Rudolf de Jong and Kees Versteegh, 309-328. Leiden: E.J. Brill.
- .-. *Forthcoming*. "Ḥassāniyya Arabic". *Encyclopedia of Arabic Language and Linguistics*, II, 000. Leiden: E.J. Brill.
- Thomason, Sarah G. and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Versteegh, Kees. 2003. "Pidginization and creolization revisited: the case of Arabic". *Approaches to Arabic dialects: A collection of articles presented to Manfred Woidich on the occasion of his sixtieth birthday*, ed. by Martine Haak, Rudolf de Jong and Kees Versteegh, 343-357. Leiden: E.J. Brill.
- Vocke, Sibylle and Wolfram Waldner. 1982. *Der Wortschatz des anatolischen Arabisch*. Erlangen: Universität Erlangen/Nürnberg.
- Wellens, Ineke. 2003. *An Arabic creole of Uganda: the Nubi language of Uganda*. Ph.D. diss., University of Nijmegen.