

## The Cross-Linguistic Function of Obligatory ‘do’-Periphrasis

ANDREAS JÄGER

The University of Queensland  
s803895@newmailbox.uq.edu.au

### 1. Introduction

The aim of this paper is to give a descriptive account of the range of functions that can be associated with obligatory ‘do’-periphrasis cross-linguistically based on a sample of 80 languages. A preliminary typological discussion of the cross-linguistic properties of this phenomenon can be found in Van der Auwera (1999). Section 2 discusses cross-linguistically identifiable criteria for ‘do’-periphrasis. In section 3 some examples are provided to show that the range of functions associated with obligatory periphrasis is limited and can be accounted for in terms of a four-way typology. Section 4 compares obligatory ‘do’-periphrasis with cases of optional use of ‘do’-periphrasis in languages outside the sample. Since languages with optional ‘do’-periphrasis often allow similar functional associations, it is proposed that the functional types argued for in this paper represent domains in which ‘do’-periphrasis is likely to become grammaticalized.

### 2. Definition of ‘do’-periphrasis

According to Crystal (1980: 262) the term “periphrasis” in a linguistic sense is the use of a combination of words for the expression of some sort of grammatical relationship instead of inflection or agglutination. The domain of inquiry in this paper will be a periphrastic verb construction minimally consisting of two elements, one being a full lexical verb and the other an auxiliary. The latter has a homophonous lexical equivalent with an abstract (or highly schematic) ‘action’ meaning as exemplified by the English word *do*. The notion of abstractness or schematicity in this respect is illustrated in table 1:

Table 1. A hierarchy of schematicity

verb	EXISTENCE	ACTION	MOTION	LEG USE	ATTITUDE
exist	+	-	-	-	-
do	+	+	-	-	-
move	+	+	+	-	-
walk	+	+	+	+	-
strut	+	+	+	+	+

“Schematic action” items such as *do* are thus to be distinguished from more specific items as well as from more abstract items like *exist*, which are likewise highly schematic, but do not contain the ‘action’ feature. While *exist* represents the highest level of abstractness/schematicity, *strut* is quite specific. Table 1 places these verbs on a hierarchy according to semantic specifications. All elements that have positive values for the first two criteria therefore count as ‘do’-elements. For cross-linguistic identification and

typological discussion of the association between form and function a potential construction in a given language has to fulfil the following criteria:

- There is a lexical/pro-verb counterpart with the conceptual specifications similar to *do* in Table 1 in the same language
- Auxiliary and lexical verb necessarily share a subject (grammatical and semantic)/in case of transitive verbs they also share the object/if there is an implied subject for the lexical verb, it is necessarily co-referent with that of the auxiliary

There is some variation with respect to the impact of the conceptual structure of ‘do’ in periphrastic constructions. In the cases examined in this paper the ‘do’-element occurs as a consequence of other grammatical factors, which are characterized in terms of their functions. There are, however, cases where the ‘do’-element itself encodes functions directly. These will not be included in the present discussion.

In its lexical use ‘do’ is transitive as well as active. In some languages the ‘do’-element is compatible with all lexical verbs regardless of their conceptual structure, as evident in the English periphrastic sentence *I do not know the answer*. In other languages only transitive and active verbs can form periphrases with ‘do’. In these languages intransitive and stative verbs generally form periphrases with ‘be’ instead and the same form-function relation obtains for both ‘do’- and ‘be’-periphrasis. The features [+ transitive] and [+ active], which are characteristic of the lexical use of ‘do’-elements, thus constrain their occurrence in periphrastic constructions in some languages. This means that there is cross-linguistic variation with respect to the degree of semantic bleaching of the ‘do’-element in periphrastic constructions of the sort investigated in this paper.

### 3. Towards a typology of obligatory ‘do’-periphrasis

While the phenomenon of ‘do’-periphrasis in English (*do*-support) has received a great deal of attention in the literature (for a diachronic account of ‘do’-periphrasis in English see (to name but a few) Ellegård (1953), Stein (1986) Kroch (1989, 1989b), Nevalainen (1991) and Kallel (2002), for a generative analysis see Pollock (1989), Chomsky (1993) and Bobaljik (1995), for a thorough review of various historical and grammatical accounts see Denison (1993)), it has not been studied extensively from a cross-linguistic perspective. Van der Auwera (1999) pioneers in making the case for a typological discussion of ‘do’-constructions, more specifically the functions associated with the ‘do’-element therein. He points out under which conditions a language is likely to use periphrasis, e.g. integration of foreign words and foreign inflections. A cross-linguistically significant between negation and periphrasis is identified and also three other frequent uses of periphrastic ‘do’ (-elements), (i) general periphrasis, (ii) accomplishment and (iii) causation. Other functions such as subordinate clause marking and aspect marking are mentioned with respect to different dialects of German and Dutch. The third use, causation, is stated as a “third meaning accessible to ‘do’-verbs” (Van der Auwera 1999: 465). My findings provide some support for the cross-linguistic form-function relations proposed by Van der Auwera, particularly for that between negative elements and ‘do’-periphrasis.

However, as the following paragraphs illustrate, there is a range of other operators that likewise trigger ‘do’-periphrasis.

Some correlations between obligatory periphrasis and certain functional domains are cross-linguistically recurrent. The following sentences exemplify the functional variation of ‘do’-periphrasis. All sentences contrast with non-periphrastic sentences in the same language, which have different grammatical properties and in which the respective function is absent, as will become clear in due course.<sup>1</sup>

(1) Dumi (Tibeto-Burman):

- a.        *an*    *im*    *-a*    *ma-*            *lita*    *-mis*    *-t*            *-a*.  
           2SG    3SG    -ERG    NEG.PERF-    tell    -do    -NONPAST    -2/3SG  
           ‘He hasn’t told you.’  
           (Van Driem 1993: 240)

Kaibab (Uto-Aztecan):

- b.        *axa-*    *nim*    *-aru* *’ai mang*  
           Q-    do    -COP    that.NOM  
           *kamunch*    *-ang*    *-sinka*    *-kain*    *-n(a)?*  
           rabbit            -NOM    -skin    -PERF    -OSP  
           ‘How did you skin that rabbit?’  
           (Bunte 1986: 294)

Cakchiquel (Mayan):

- c.        *š*                    *in-*                    *b’an*    *pensar*.  
           COMPL            1SG.ERG            do    think  
           ‘I thought.’  
           (Stenson 1998: 224)

Korku (Munda):

- d.        *ra:mra:m*    *men*    *-do*    *-ne*    *bijj*    *-en*.  
           ramram    say    -do    -ADV    get.up    -PERF  
           ‘Got up saying „ramram”.’  
           (Nagaraja 1999: 79)

In the first example a ‘do’-element becomes obligatory as soon as a closed class morpheme is present in the clause, in this case a negative morpheme. The sentence therefore contrasts with a non-periphrastic counterpart that lacks this morpheme. This sentence exemplifies a grammatical environment in which obligatory periphrasis commonly occurs. In Dumi sentences like (1a) are construed periphrastically only when the lexical verb carries the

<sup>1</sup> Abbreviations used in this paper are: ACC, accusative; ADV, adverbializer; AGR, agreement; ASP, aspect; AUG, augmented; AUX, auxiliary; COMPL, complementizer; CONT, continuous; COP, copula; DET, determiner; DIST, distal deictic aspect; ERG, ergative; FACT, factual; FOC, focus; INES, inessive; INF, infinitive; INT, intensive; MAR, Madang-Adelbert Range; NOM, nominative; OSP, oblique subject particle; PERF, perfect; PL, plural; PRED, predicate; PRES, present; PROG, progressive; PRT, particle; PTC, participle; PURP, purposive; Q, interrogative; REFL, reflexive; SG, singular; SS, same subject; TOP, topic; TNG, Trans New-Guinea phylum; VN, verbal noun.

negative prefix *ma-*, which also encodes perfectivity. If this is absent, other regular verb morphology is realized on the lexical verb and not on the ‘do’-form. In all languages of the sample the trigger elements for ‘do’-periphrasis are from a closed class. The functional range associated with these elements is limited and it is widely congruent with those associated with verbal categories commonly encoded in inflections. This type of periphrasis is to be differentiated from cases where the occurrence of a ‘do’-element directly changes the meaning of its non-periphrastic counterpart, independent of other material.

Example (1b) shows a deviation from the canonical word order that obtains in the language, i.e. ‘do’-periphrasis occurs only in clauses that have non-canonical word-order. The function usually associated with the non-canonical word order in periphrastic sentences shows considerable cross-linguistic similarities. The deviation is typically associated with interrogative or topic/focus function, an observation that is borne out by this example. In Kaibab periphrasis is not triggered by additional morphological material in the clause. Here a ‘do’-element is used as a carrier for the bound interrogative morpheme and the entire complex occurs clause-initially.

Example (1c) shows a tighter connection between the lexical verb and ‘do’ than (1a) and (1b). There are constraints on the compatibility of certain verb stems with inflectional morphology. In Cakchiquel most verb stems inflect. There are, however, certain stems that disallow the direct affixation of verbal morphology. In these cases ‘do’-periphrasis is obligatory. Its occurrence depends on inherent lexical features of the verb. The feature [+borrowed] such as in the Cakchiquel example is a common trigger for ‘do’-periphrasis cross-linguistically. The function of ‘do’-periphrasis here can be described as analytic verb-/conjugation class marking.

The last example shows ‘do’-periphrasis as a means to mark a verbal constituent as adverbial or converbal. As such the thus marked lexical verb is assigned the function of secondary predication subordinated to a primary predicator.

My investigation has shown that wherever ‘do’-periphrasis is obligatory in certain clause-types in a given language, one of the functional associations that obtain in (1) can be identified. In some cases a combination thereof can be observed, i.e. there are languages that employ periphrasis under various conditions, but nevertheless this variation is limited to the conditions exemplified by (1). Consequently I propose that cross-linguistically the functional range of ‘do’-periphrasis in environments where it is obligatory is limited and can be characterized in terms of the following four types:

- Type 1     It is the appearance of lexical or morphological material that triggers ‘do’-periphrasis, in most cases material that attaches to the lexical verb and thus prevents the realization of regular verbal grammatical categories as affixes on the verb. This material usually belongs to a closed class and its function is similar to that of regular verbal categories and/or adverbial modification.
- Type 2     If a language has rigid or dominant word order, periphrasis is used to mark clause types that display a deviant or irregular word order or to maintain a close approximation of the regular word order in these, i.e. keeping the

relative order of verb and object unchanged. Focalization, topicalization and interrogativity are the most common (pragmatic) functions that can be associated with periphrasis of this type.

- Type 3 Lexical features of the verb require periphrasis with a ‘do’-auxiliary, which is typically either inherent native features or the fact that the lexical verb is a loanword.
- Type 4 If a verb or verb phrase forms a constituent in a larger structure, such as converbial structures, it is obligatorily marked as such by means of ‘do’ - periphrasis.

In the following paragraphs I will provide further language data to support this claim and to make a more fine-grained typological description possible. Languages often display more than one of these types of periphrasis. For instance if a language uses ‘do’-periphrasis obligatorily with a certain class of verbs, other factors such as word order considerations or trigger morphemes may then facilitate a clearer separation of verb and ‘do’-auxiliary rather than being directly responsible for periphrasis. Morphological trigger material may facilitate a divergence from canonical word order and consequently periphrasis.

As stated above, it is a cross-linguistic tendency for certain morphemes to disallow co-occurrence with inflected lexical verbs. Periphrasis with a ‘do’-auxiliary provides a carrier for these inflections. My investigation has shown that in all languages that use this type of periphrasis the trigger morphemes are members of a closed class and encode either clausal negation (examples (1a) and (2a-c)) or modify the meaning of the lexical verb (examples (2d-h)).

(2) Nanai (Tungusic):

- a.      əm      xolā      ta      -xan  
           NEG    read    do      -PAST  
           ‘did not read’  
           (Payne 1985: 230)

Lezgian (Caucasian):

- |    |                |                 |                |
|----|----------------|-----------------|----------------|
|    | AFFIRMATIVE:   | NEGATIVE:       |                |
| b. | <i>awun</i>    | <i>t-</i>       | <i>awun</i>    |
|    | do (lexical V) | NEG-            | do             |
| c. | <i>čüxün</i>   | <i>čüxün t-</i> | <i>awun</i>    |
|    | wash           | wash NEG-       | do (auxiliary) |
|    |                | <i>*t-</i>      | <i>čüxün</i>   |
|    |                | NEG-            | wash           |
- (Haspelmath 2000: 658)

Carrier (Athabaskan):

- d.     nʌlgaih        za       ʔʌtʼen.  
        3SG.run       only   3SG.do  
        ‘He kept on running.’
- e.     xwʌninʔai     dizkʼan        za       ʔʌtʼen.  
        pole           3SG.burn       only   3SG.do  
        ‘The pole keeps on burning.’  
        (Poser 2002: 4-7)

Imonda (TNG/Northern):

- f.     uòs    lèha   -kubui fe   -f.  
        Moon shine -INT do   -PRES  
        ‘The moon shines brightly.’  
        (Seiler 1985: 111)

Toura (Niger-Congo):

- g.     tià    -ʔ    gwɛɛ        lɔ    -ʔ    woʔ    le.  
        Tia    -PRED peanuts    buy   -FOC do    PAST  
        ‘Tia **bought** peanuts.’
- h.     tià    -ʔ    gwɛɛ        -ʔ    lɔʔ.  
        Tia    -PRED peanuts    -FOC buy  
        ‘Tia bought **peanuts**.’  
        (Van Valin 1999: 511-24)

In (2d-e) the periphrastic construction yields a progressive reading, a function that is commonly encoded by the ‘do’-auxiliary itself across languages. In Carrier, however, the element *za* (‘only’) triggers periphrasis. Imonda likewise uses regular verb inflection in declarative clauses unless the lexical verb carries an affix from a closed class, such as *-kubui* in (2f), in which case regular verb morphology is realized on a following ‘do’-form. Negative elements could also be identified as a trigger for ‘do’-periphrasis in Dumi, English, Imonda, Korafe, Korean, Kunimaipa, Maisin, Menya, Old Japanese, Persian, Siroi, Suena and Yareba. In Ainu, Cavineña, Imonda, Jacaltec, Japanese, Nabak, Selepet, Skou, and Tamil the trigger morphemes encode verbal categories such as tense, aspect and mood, while adverbial modification could be identified as their function in Enga, Jacaltec, Kashmiri, Korean and Maisin.<sup>2</sup>

In all languages with Type 1 periphrasis the trigger material encodes regular verbal categories, where these are not realized as inflections on the lexical verb, or adverbial modification. No languages/examples could be found where material functioning in other domains triggers obligatory verb periphrasis with a ‘schematic action’-auxiliary. If the closed class of trigger morphemes contains only one item, this can be characterized as a clausal negator morpheme in most cases.

<sup>2</sup> See table 2 for genetic affiliations.

As stated above, deviations from non-canonical word order commonly result in topic, focus or interrogative readings. In a number of languages these clause types require a ‘do’-auxiliary to carry all or some inflections. Consider the following examples:

(3) Skou (Sko/Vanimó):

a. *báng moeritó ke- k- ang.*  
 yesterday yellowtail scad (y.s.) 3SG- 3SG- eat  
 ‘He ate some yellowtail scad yesterday.’

b. *moeritó ke- k- ang -inga báng ke- li.*  
 y.s. 3SG- 3SG- eat -the yesterday 3SG- do  
 ‘Eat yellowtail scad he did yesterday.’  
 (Donohue 2003: 109)

Upriver Halkomelem (Salishan):

c. *lí -cxw yáyes?*  
 do -2SG work.PROG  
 ‘Did you work?’

d. *lí -cε spēpek?*  
 do -FUT float  
 ‘Will it float?’  
 (Galloway 1993: 355)

Tarifit (Berber):

e. *i -zra a -Kidar.*  
 3SG -see 3SG -Kidar  
 ‘He saw the horse.’

f. *wi -g -izri -n -a -Kidar?*  
 who -do -see -PTC -3SG -Kidar  
 ‘Who saw the horse.’<sup>3</sup>

Apalai (Arawakan):

g. *oty ka -se yto -vko m- a?*  
 Q do -PURP go -CONT 2 be.PRES  
 ‘Why are you going?’  
 (Koehn & Koehn 1986: 44)

Yoruba (Niger-Congo):

h. *bàwo ni še lí fẹ̀ iyàwó ní ilè Yorùbá?*  
 how be.3SG do ASP marry wife be ground Yoruba  
 ‘How does one get married in Yorubaland?’  
 (Yusuf 1992: 117)

<sup>3</sup> Abdel El-Hankari (personal communication)

Similar observations could also be made in Ainu, Apalai, Bariba, Colonial Yucatec, English, Eseejja, Guanano, Hausa, Ivie, Kaibab, Maricopa, Mojave, Paez, Paumarí, Rukai, Skou, Squamish, Tarifit, Tauya, Ute, Yimas and Yoruba.<sup>4</sup> In topicalization structures it is always the verb or verbal predicate that is topicalized periphrastically, while other topicalized constituents do not trigger periphrasis. There is a cross-linguistic tendency to associate ‘do’-periphrasis functionally with focalization and topicalization on the one hand, and interrogativity on the other hand. Formally the former goes hand in hand with mostly non-canonical clause-initial position of the focalized or topicalized constituent. Three functional subtypes can be distinguished here, based on what is focalized or topicalized: 1. Predicate, 2. Argument or property of argument, 3. Whole clause. If interrogativity is involved, it is manner interrogatives rather than polar interrogatives in the majority of languages that require a ‘do’-auxiliary.

The data in (4) exemplify some instances of periphrasis triggered by features of the lexical verb stem. In most languages in the sample only a subclass of verb stems require obligatory ‘do’-periphrasis, while other verb stems carry inflections.

(4) Pipil (Uto-Aztecan):

- a. *yah nu- chiw -ki arrepentir.*  
 3SG REFL- do -PAST repent  
 ‘They regretted (it).’  
 (Campbell 1985: 144)

Bangla (Indo-Iranian):

- b. *mægnifai kōraa*  
 magnify do  
 ‘to magnify’  
 (Bhattacharya 2001: 15)

Ungarinjin (Wororan):

- c. *wulan ŋ- ama -naŋga.*  
 speak 1SG- do -to.him  
 ‘I speak to him.’  
 (Capell 1979: 233)

Basque (Isolate):

- d. *nor -k egin zuen negar?*  
 who -ERG do AUX cry  
 ‘Who cried?’  
 (Ortiz de Urbina 1989: 44)

Ingush (Caucasian):

- e. *geanaž kag j- i -e.*  
 branch.PL break AGR- do -INF  
 ‘to break branches’  
 (Nichols 1994: 112)

<sup>4</sup> See table 2 for genetic affiliations.



The periphrastically marked subclass is commonly defined in terms of the feature [+borrowed]. However, it can also comprise native verb stems, such as in Basque or Ingush.

Note in (4d) that although lexical verb stem and ‘do’-auxiliary are mostly adjacent, in some contexts this constraint does not apply. These contexts generally correspond to functional domains in which periphrasis often occurs in other languages, such as interrogativity. Lexically conditioned ‘do’-periphrasis could be identified in the following languages: Basque, Budukh, Cakchiquel, Chechen, Erromangan, Hindi, Imonda, Ingush, Jingulu, Kannada, Meithei, Montagnais, Persian, Rama, Rutul, Suená, Sumu, Turkish, Wardaman and, once more, Yoruba.<sup>5</sup>

The data in the sample shows that lexically conditioned periphrasis comes in three varieties: 1. ‘Do’-periphrasis marks a subclass of verbs, which is not necessarily semantically coherent. The auxiliary is needed to form a finite verb/predicate and is comparable in function to conjugation class marking. No cross-linguistically valid semantic characterization of such a verb class could be identified. 2. ‘Do’-periphrasis integrates loanverbs in target languages that disallow the direct combination with native verb morphology. 3. A ‘do’-element is the default auxiliary in languages that have a large stock of unmarked verb stems and realize all morphology on an auxiliary. The subclass of verbs in such languages equals the class of verbs as a whole and periphrasis is the standard strategy for verbal predication. That is to say each lexical verb is basically an unmarked stem which requires an auxiliary to appear in a grammatical clause.

The following examples have in common that ‘do’-periphrasis obligatorily marks the subordination of a constituent. Often it is two symmetrically coordinated lexical verbs that are marked periphrastically as a constituent, such as in (5b,c and e).

(5) Tshangla (Tibeto-Burman):

- a. *jang ma- ke -la -n chho -la -gai*  
 1SG NEG- be.born-PTC -do stay -PTC -ABL  
*goma ama mongshi thong -ma.*  
 before mother dream see -PAST  
 ‘Before I was born, my mother had a dream.’  
 (Andvik 2004: 326)

Enga (TNG/Main Section):

- b. *akáli dokó -mé mená dupá itáki*  
 man DET -AG pig DET.PL count  
*pyá -pa kayá -pa pyó -o pi*  
 hit -COMPL stop -COMPL do -O hit  
*-ly -á -mo.*  
 -PRES -3SG -AUG  
 ‘That man is counting the pigs over and over again.’  
 (Lang 1973: xlv)

<sup>5</sup> See table 2 for genetic affiliations.

Thakali (Tibeto-Burman):

- c. *t'am -cä cä -pa cä -pa la -si mu.*  
 3PL -PL look -VN look -VN do -CONVERB COP  
 'They look at each other.'  
 (Georg 1995 138)

Cubeo (Tucanoan):

- d. *nocacu ui pupui -buchibure*  
 there.one his bewitched -cigar  
*nurĩ -teyu baacu edamede.*  
 smoke -do who.be.PAST come.3.PAST  
 'He came, having smoked the bewitched cigar of the one who lives there.'  
 (Salser & Salser 1977: 267)

Chantyal (Tibeto-Burman):

- e. *fəya bəw fəya ama la -gəy*  
 oh father oh mother do -PROG  
*khuy khuy la -gəy*  
 pant pant do -PROG  
*la -wa -nfiari tuŋ pfiə -i tə.*  
 do -NOM -INES up take.away -PERF FACT  
 'I took it up while going "Oh father, oh mother" and panting.'  
 (Noonan 1999: 5)

This type of periphrasis is the least common cross-linguistically. Nevertheless, as the data shows, a considerable number of languages makes use of periphrasis for this function. Apart from those in (5) the following languages allow this form-function relation: Chantyal, Desano, Fore, Kobon, Korku, Meithei, Siane and Tarangan.<sup>6</sup>

The form-function relations proposed for Types 1 and 2 show some similarity. For instance, topicalization of the predicate is achieved by affixation of a trigger morpheme in one language, while in others it requires a word order change, generally speaking the fronting of the predicate. In both cases 'do'-periphrasis is obligatory as a consequence of one of these circumstances. Generative analyses of obligatory *do*-support in English (see for instance Pollock 1989, Bobaljik 1995) have pointed out the syntactic correlation between word order change in interrogatives on the one hand and operators such as clausal negators on the other hand. A common grammatical explanation underlying these types is therefore not unlikely also from a cross-linguistic perspective. In Type 3 ease of processing may be held responsible for periphrasis with a 'do'-auxiliary. The use of an inflected auxiliary in a large number of verb constructions provides an avoidance strategy for morphological complexity or uncertainty in the case of borrowed stems, since the speaker only has to acquire a single paradigm, namely that of the 'do'-auxiliary. In fact a tendency to employ 'do'-periphrasis optionally has been observed in (first) language acquisition for instance in Dutch and English (Tieken-Boon van Ostade 1990: 20-21, Van der Auwera 1999: 60). While a detailed explanation is beyond the scope of the present typological study, possible correlations between the types are a fertile ground for future research.

<sup>6</sup> See table 2 for genetic affiliations.

The distribution of functions associated with obligatory ‘do’-periphrasis is summarized in table 2. Some languages show ‘do’-periphrasis of more than one type. No languages could be found in the sample that employ ‘do’-periphrasis obligatorily for other functions than those accounted for by the four-way typology advocated in this paper.

**Table 2.** Functions associated with obligatory ‘do’-periphrasis

Language	Genetic affiliation	Type 1	Type 2	Type 3	Type 4
Ainu	Isolate	+	+		
Apalai	Macro-Carib		+		
Bangla	Indo-Aryan			+	
Bariba	Niger-Congo		+		
Basque	Isolate		+	+	
Budukh	Caucasian			+	
Cakchiquel	Mayan			+	
Carrier	Athabaskan	+			
Cavineña	Ge-Pano	+	+		
Chantyal	Tibeto-Burman				+
Chechen	Caucasian			+	
Col.Yucatec	Penutian		+		
Cubeo	Tucanoan				+
Desano	Tucanoan				+
Dumi	Tibeto-Burman	+			
Enga	TNG/Main Section	+			
English	Germanic	+	+	+	
Erromangan	Malayo-Polynesian			+	
Eseejja	Tacanan		+		
Fore	TNG/Main Section				+
Guanano	Tucanoan		+		
Halkomelem (Upriver)	Salishan		+		
Hausa	Chadic		+		
Hindi	Indo-Iranian			+	
Hualapai	Hokan			+	
Imonda	TNG/Northern	+		+	
Ingush	Caucasian			+	
Ivie	Niger-Congo		+		
Jacaltepec	Mayan	+	+		
Japanese	Altaic	+			
Kaibab	Uto-Aztecan		+		
Kanite	TNG/Main Section	+			
Kannada	Dravidian			+	
Kashmiri	Indo-Iranian	+			
Kobon	TNG/Main Section				+
Korafe	TNG/Main Section	+			
Korean	Altaic	+			
Korku	Munda				+
Kunimaipa	TNG/Main Section	+			
Lezgian	Caucasian	+			
Lori	Indo-Iranian			+	
Maisin	Malayo-Polynesian	+			
Maricopa	Hokan	+	+		
Meithei	Tibeto-Burman				+
Menya	TNG/Main Section	+			
Mojave	Hokan		+		
Montagnais	Algonquian			+	
Nabak	Papuan	+			
Nanai	Tungusic	+			
Old Japanese	Altaic	+			
Paez	Chibchan-Paezan		+		
Paumarí	Arauan		+		

Language	Genetic affiliation	Type 1	Type 2	Type 3	Type 4
Persian	Indo-Iranian	+		+	
Pipil	Uto-Aztecan			+	
Rama	Chibchan-Paezan			+	
Rukai	Formosan		+		
Rutul	Caucasian			+	
Selepet	TNG/Main Section	+			
Siane	TNG/Main Section				+
Siroi	TNG/MAR	+			
Skou	Sko/Vanimó	+	+		
Squamish	Salishan	+	+		
Suena	TNG/Main Section	+			
Sumu (Northern)	Chibchan-Paezan			+	
Tarifit	Berber		+		
Tamil	Dravidian	+			
Tarangan	Malayo-Polynesian				+
Tariana	Tucanoan		+		
Tauya	TNG/Main Section		+		
Thakali	Tibeto-Burman				+
Toura	Niger-Congo	+			
Tshangla	Tibeto-Burman				+
Turkish	Turkic			+	
Ungarinjin	Wororan			+	
Ute	Uto-Aztecan		+		
Wardaman	Gunwingguan			+	
Yangman	Gunwingguan			+	
Yareba	TNG/Main Section	+			
Yimas	Sepik-Ramu		+		
Yoruba	Niger-Congo		+	+	

#### 4. Optional ‘do’-periphrasis

A correlation between the use of periphrasis and non-syntactic factors can be observed cross-linguistically. Wherever there is a strong correlation between pragmatic function and periphrasis, there is some degree of optionality. Periphrasis can be triggered by several or a combination of the grammatical factors mentioned in section 3. For instance, a clause type that requires obligatory periphrasis may be marked by and consequently defined in terms of the presence of morphological material. If morphological trigger material is functionally limited, it can be expected that also the clause types that are environments for periphrasis have a limited functional range cross-linguistically. Moreover, since the conceptual content of the ‘do’-element has more or less strong influence on the use in obligatory periphrasis and strict obligatoriness does not always obtain, the same ‘do’-element can reveal various levels of grammaticalization within one and the same language.

Once the functional range has been described in contexts where it is obligatory, it can serve as the basis for predictions on future grammaticalization. There are functional similarities between obligatory and optional periphrasis. That is to say that language A employs ‘do’-periphrasis optionally in contexts that are functionally similar to those contexts that make it obligatory in language B.

If a language employs optional ‘do’-periphrasis it means that the speaker can express the same propositional content either non-periphrastically or periphrastically. Languages that employ ‘do’-periphrasis differ in terms of whether they use it exclusively for a given function or whether they use it as one of several strategies for the same function. The

functional difference lies in the pragmatic context in which these options are likely choices. Some examples of optional periphrasis are given below:

(6) German (Germanic):

- a. *Ich arbeite.*  
 1SG work.1SG.TOP  
*Ich tue arbeiten.*  
 1SG do.1SG work.INF  
 'I work.'

Ika (Chibchan):

- b. *zoža u -na.*  
 go do -DIST  
 'They went..'

- c. *mouga -ri awa?rei zoža -na.*  
 two -TOP below go -DIST  
 'Two men went below.'  
 (Frank 1990: 48-49)

In both languages 'do'-periphrasis is optional in declarative clauses. In German the periphrastic option can be chosen to highlight or topicalize the action encoded in the lexical verb and contrast it with other actions.<sup>7</sup> More specifically, the verb is fronted and a periphrastic auxiliary takes its place. Without a word order change intonation fulfils a similar function. According to Frank (1990) Ika optionally employs the periphrastic variant to introduce a new verb to the discourse. In a number of the languages discussed in section 3 above 'do'-periphrasis is the standard, if not the only strategy for these functions.

Periphrastic constructions occur optionally also in Nisenan. Here the auxiliary carries agreement morphology, while the main verb remains non-finite. Switch-reference morphology, which otherwise fulfils its expected function in the language, also turns up on auxiliary 'do', but here its function is suspended, which suggests verb periphrasis. That is to say that even if one element is marked with a 'different subject'-morpheme, there is only one subject in the structure.

(7) Nisenan (Penutian):

- a. *aanik'in kas myi -i mee -naa -i?*  
 want.SS do.1SG it -ACC get -CONSEQUENTIAL -INF  
 'If I wanted it, wouldn't I get it?'  
 (Eatough 1999: 14)

<sup>7</sup> Periphrasis with 'do' is marked as colloquial in Standard German, but widely used in dialectal varieties.

Breton (Celtic):

b. \*bremañ e welan e  
 now PRT see.1SG PRT  
*ra len Lenaig al levr.*  
 do.3SG read Lenaig the book  
 ‘I see now that Lenaig reads the book.’

c. *bremañ e welan he deus lennet*  
 now PRT see.1SG 3SG have.3SG read  
*Lenaig al levr.*  
 Lenaig the book  
 ‘I see now that Lenaig has read the book.’  
 (Stephens 1993: 395)

Periphrastic constructions in Nisenan are subject to the constraint that they are only permitted in matrix clauses. Since they are optional even here, some further pragmatic function is likely. The matrix clause constraint also goes for Breton, as the ungrammaticality of (7b) shows, but nevertheless it is optional even in matrix clauses. No languages could be found that employ periphrasis obligatorily for the marking of matrix clauses.

Optional ‘do’-periphrasis quite commonly operates within the same functional domains as obligatory periphrasis. This suggests that there are certain functional domains, in which the construction is more likely to be grammaticalized in than in others. Obligatory periphrasis can be seen as a grammaticalized form-function relation, which entails that the examples in (6) are more likely to be grammaticalized than those in (7).

## 5. Concluding remarks

If a language employs ‘do’-periphrasis obligatorily, the resulting form-function relation can be characterized only in terms of either one of the four types stated in section 2 or a combination thereof. If a language employs ‘do’-periphrasis optionally and a similar form-function relation can be established, this form-function relation is likely to be grammaticalized, i.e. obligatorily expressed by means of ‘do’-periphrasis. If, on the other hand, a different form-function relation can be established, grammaticalization is not imminent, since there are no attested cases of obligatory periphrasis, in which the same form-function relation obtains.

## References

- Andvik E 2004 ‘“Do” as subordinator in Tshangla’ in A Saxena (ed) *Himalayan linguistics - past and present* Mouton de Gruyter Berlin: 311-340.  
 Bhattacharya T 2001 ‘Bengali’ in J Garry & C Rubino (eds) *Facts about the world’s languages: an encyclopedia of the world’s major languages: past and present* H.W. Wilson Press New York: available from  
<http://www.homepages.ucl.ac.uk/~uclyara/list.htm> (accessed 06/04)  
 Bobaljik JD 1995 *Morphosyntax: the syntax of verbal inflections* PhD dissertation MIT.

- Bunte PA 1986 'Subordinate clauses in Southern Paiute' *International Journal of American Linguistics* 5: 275-300.
- Campbell L 1985 *The Pipil language of El Salvador* Mouton de Gruyter Berlin.
- Capell A 1979 'Classification of verbs in Australian languages' in S Wurm (ed) *Australian linguistic studies* Pacific Linguistics Canberra: 259-323.
- Chomsky N 1993 *The Minimalist program* MIT Press Cambridge MA.
- Crystal D 1980 *A first dictionary of linguistics and phonetics* André Deutsch London.
- Denison D 1993 *English historical syntax: verbal constructions* Longman London.
- Donohue M 2003 *A grammar of the Skou language of Papua New Guinea* Ms University of Sydney: available from <http://courses.nus.edu.sg/course/ellmd/Skou/> (accessed 11/03)
- Eatough A 1999 *Central Hill Nisenan texts with grammatical sketch* University of California Press Berkeley.
- Ellegård A 1953 *The auxiliary DO: the establishment and regulation of its use in English* Stockholm Almqvist & Wiksell.
- Frank P 1990 *Ika syntax: studies in the languages of Colombia I* Summer Institute of Linguistics & The University of Texas Dallas.
- Galloway BD 1993 *A grammar of Upriver Halkomelem* University of California Press Berkeley.
- Georg S 1995 *Marphatan Thakali - Untersuchungen zur Sprache des Dorfes Marpha im oberen Kāli- Gandali-Tal, Nepal* Lincom Europa München.
- Haspelmath M 2000 'Periphrasis' in B Booij, C Lehmann, J Mugdan (eds) *Morphologie - ein internationales Handbuch zur Flexion und Wortbildung, 1. Halbband* Walter de Gruyter Berlin: 654-664.
- Kallel A 2002 'The age variable in the rise of periphrastic 'do' in English' *Reading Working Papers in Linguistics* 6: 161-185.
- Koehn E & Koehn S 1986 'Apalai' in DC Derbyshire & GK Pullum (eds) *Handbook of Amazonian languages, Volume I* Mouton de Gruyter Berlin: 33-127.
- Kroch A 1989 'Function and grammar in the history of English: periphrastic 'do'' in FW Fasold & D Schiffrin (eds) *Language Variation and Change: current issues in linguistic theory, vol. 52* John Benjamins Amsterdam: 133-172.
- Kroch A 1989b 'Reflexes of grammar in patterns of language change' *Language variation and change 1*: 199-244.
- Lang A 1973 *Enga Dictionary* Pacific Linguistics Canberra.
- Nagaraja KS 1999 *Korku language grammar, texts and vocabulary* Tokyo Institute for the study of languages and cultures of Asia and Africa.
- Nevalainen T 1991 'Motivated archaism: the use of affirmative periphrastic *do* in Early Middle English liturgical prose' in D Kastovsky et al. (eds) *Historical syntax (Topics in English linguistics 2)* Mouton de Gruyter Berlin: 303-320.
- Nichols J 1994 'Ingush' in R Sweets (ed) *The indigenous languages of the Caucasus, Vol., North East Caucasian languages, Part 2* Caravan books Delmar: 77-145.
- Noonan M 1999 'Converbal constructions in Chantyal' in P Yadava & WW Glover (eds) *Topics in Nepalese Linguistics* Royal Nepalese Academy Kathmandu: available from <http://www.uwm.edu/~noonan/Papers.html> (accessed 08/03)

- Ortiz de Urbina J 1989 *Parameters in the grammar of Basque: A Government and Binding approach to Basque syntax* Foris Publications Dordrecht.
- Payne J 1985 'Negation' in T Shopen (ed) *Language typology and syntactic description, Vol. 1* Cambridge University Press Cambridge: 197-242.
- Pollock JY 1989 'Verb movement, universal grammar and the structure of IP' *Linguistic Inquiry* 20: 365-424.
- Poser W 1999 'Scope and dummy verbs in Carrier' *University of British Columbia Working Papers in Linguistics* 2: available from <http://www.cis.upenn.edu/~wjposer/.downloads/sdvc.pdf> (accessed 10/02)
- Salser, J.K. & Salser, N. 1977 'Some features of Cubeo discourse and sentence structure' in RE Longacre & F Woods (eds) *Discourse grammar: studies in indigenous languages of Colombia, Panama, and Ecuador, part 2* Summer Institute of Linguistics and the University of Texas at Arlington Dallas: 253-272.
- Seiler W 1985 *Imonda, a Papuan language* Pacific Linguistics Canberra.
- Stein D 1986 'Syntactic variation and change: the case of DO in questions in Early Modern English' *Folia Linguistica Historica* 7:121-149.
- Stenson N 1998 'Spanish loanwords in Cakchiquel' in L Hinton & P Munro (eds) *Studies in American Indian languages - description and theory* University of Berkeley Press Berkeley: 223-233.
- Stephens, J. 1993 'Breton' in M Ball & J Fife (eds) *The Celtic languages* Routledge London: 349-409.
- Tieken-Boon van Ostade I 1990 'The origin and development of periphrastic do: a case of destigmatisation' *North Western Europe Language Evolution (NOWELE)*16: 3-52.
- Van der Auwera J 1999 'Periphrastic 'do': typological prolegomena' in GAJ Tops (ed) *Thinking English grammar: to honour Xavier Dekeyser, professor emeritus* Peeters Leuven.
- Van Driem G 1993 *A grammar of Dumi* Mouton de Gruyter Berlin.
- Van Valin RD 1999 'A typology of the interaction of focus structure and syntax' in E Raxilina & J Testelec (eds) *Typology and the theory of language: from description to explanation* Languages of Russian culture Moscow: available from [http://linguistics.buffalo.edu/research/rrg/vanvalin\\_papers/focuswotypology.pdf](http://linguistics.buffalo.edu/research/rrg/vanvalin_papers/focuswotypology.pdf) (accessed 05/04)
- Yusuf O 1992 'Variant orders and the discourse use of bleached verbs in Yoruba' *Afrikanistische Arbeitspapiere* 30: 113-134.