Secondary Publication

Haig, Geoffrey

Noun-plus-verb complex predicates in Kurmanjî Kurdish : Argument sharing, argument incorporation, or what?

Date of secondary publication: 20.12.2023

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-925144

Primary publication

Haig, Geoffrey (2002): „Noun-plus-verb complex predicates in Kurmanjî Kurdish : Argument sharing, argument incorporation, or what?“. In: Language typology and universals = Sprachtypologie und Universalienforschung : STUF, Vol. 55, Nr. 1, pp. 15-48, Berlin: De Gruyter Mouton, doi: 10.1524/stuf.2002.55.1.15.

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holder(s).

This document is made available with all rights reserved.

Sprachtypol. Univ. Forsch. (STUF), Berlin 55 (2002) 1, 15–48

GEOFFREY HAIG (Kiel)

Noun-plus-verb complex predicates in Kurmanjî Kurdish: Argument sharing, argument incorporation, or what?[1]

Abstract

Complex predicates (CPs) consisting of a noun (N) and a verb (V) are an ubiquitous feature of Kurdish, and of Indo-Iranian and Indo-Aryan languages generally. MOHANAN (1997) has proposed an argument-sharing analysis for this type of CP in Hindi, according to which both the noun and the verb contribute to the argument structure of the CP. In this paper the argument-sharing approach is assessed against the Kurdish data, but it transpires that it only accounts for a subset of N+V CPs. Furthermore, for one specific type of CP, an analysis in terms of syntactic incorporation is simpler and empirically more adequate. I conclude that no single model accounts for the totality of CP-formation in Kurdish and related languages. Finally, I address the question of why N+V CPs should have emerged in the eastern members of Indo-European, yet are almost completely lacking in the Indo-European languages of Europe.

1. Introduction

Complex predicates (CPs) have attracted considerable attention in recent linguistics (see e.g. CATTELL 1984, GRIMSHAW & MESTER 1988, the contributions to ALSINA et al. 1997 and SCHULTZE-BERNDT 2000). In this paper, I will focus on one particular type of CP, namely those consisting of a noun (N) and a verb (V). An example of this type of CP from Kurdish is *maç kirin* 'kiss do/put':

(1) *Ferho dest-ê mel-ê* ***maç*** ***kir*** *[...]* [MEY 43]
Ferho hand-L.M mollah-OBL kiss do:PST(3S)
'Ferho kissed the hand of the mollah' (lit. 'kiss-put/did the hand of ...')

The fundamental linguistic problem with CPs of this type is that while they are composed of two recognizable and distinct lexical items, they behave in many respects like a single unit. As far as the syntactic analysis of CPs are concerned, two related issues are at stake: First, the precise status of the N: is it an argument of the V, or has it been entirely incorporated into the verbal complex, or does it possess some kind of intermediate status? Second, how do such composite predicates determine the structures of the clauses they head, and in particular, just what does the N contribute to the argument structure?

[1] This research was conducted as part of the project *Grammatische Relationen im Kurmanjî-Kurdisch*, funded by the Deutsche Forschungsgemeinschaft, Grant no. Mo728/2–1. I am very grateful to the project supervisor, ULRIKE MOSEL, for her support throughout, and to NICOLE NAU for detailed comments on earlier versions of this paper. The responsibility for the remaining errors is entirely my own.

In this paper I will be proposing some answers to these and related questions for N+V CPs in Kurdish, and relating them to existing proposals for comparable structures in other languages. Two alternatives are considered. The first is the argument-sharing approach, spelled out for CPs in Hindi by Mohanan 1997. On her view, the N is an argument-assigner, and the N and the V jointly determine the argument structure of the clause. An alternative view sees the N as incorporated into the verbal complex, i.e. the N is not an argument assigner (Rosen 1989). To anticipate my conclusions, neither approach satisfactorily accounts for the full range of data in Kurdish. Instead, each provides a plausible explanation for one particular type of CP, but fails to provide a convincing account of CPs as a particular typological feature of the language.

The paper is organized as follows: In § 2 I look briefly at CPs from a typological perspective. Section 3 gives background information on Kurdish, and provides a brief sketch of the relevant features of the verb system. In § 4, my analysis of CPs in Kurdish is presented, and in § 5 I evaluate the Kurdish data against the argument-sharing account of Hindi CPs in Mohanan (1997). In § 6, a number of residual issues are treated, and some suggestions are made on the diachronic emergence of CPs in Kurdish and related languages.

2. Complex predicates from a cross-language perspective

Complex predicates are informally defined by Alsina et al. 1997 as "predicates which are multi-headed; they are composed of more than one grammatical element (either morphemes or words), each of which contributes part of the information normally associated with a head". Yet even a definition as broad as this one fails to cover all the attested uses of the term "complex predicate", as a glance through the contributions to the volume from which the definition is taken reveals. The following phenomena are all treated as "complex predicates", whereby this list is by no means exhaustive:

1. Morphologically simplex unergative verbs in English such as *dance* (Hale & Keyser 1997)
2. Combinations of "particles" plus a verb lexeme, e.g. Hungarian *reki rohan*, lit. 'toward rush' (Ackerman & Lesourd 1997)
3. Predicates consisting of a single morphologically complex word, e.g. causatives in Chichewa (Alsina 1997)
4. Serial verb constructions containing two or more verb lexemes (Durie 1997)
5. Combinations of auxiliary and main verb in some languages of Europe (Rosen 1997)
6. Combinations of a noun plus a verb, e.g. in Hindi (Mohanan 1997)

Although the drive to discover uniformity across diversity is understandable, the current extension of the term "complex predicate" makes it difficult to delineate a precise area of cross-language investigation. For example, once we relax our definition to include morphologically simple predicates, as advocated by Hale & Keyser 1997, lexical causatives such as English *kill* fall within the range of complex predicates, to name but one of many potential candidates. While this is in fact the intended result of the approach of Hale & Keyser, the difficulties from a cross-language perspective should be evident (and even the justification from an English perspective is arguable, see Kiparsky 1997). The conclusion I draw from all this is that if the term "complex predicate" is really intended to cover all the phenomena listed above, then it does not constitute a viable domain for cross-linguistic comparison.

The present paper is therefore limited to a sub-section of complex verbs, namely CPs consisting of a noun and a verb, i.e. option 6 above. Many languages have N+V CPs of this type. In Indo-Aryan languages, they are extremely widespread. Traditionally they are termed "conjunct verbs" (MASICA 1991: 368) to distinguish them from "compound verbs", sequences of two or more verbs (see BUTT 1997 for recent treatment of V+V complex verbs in Hindi/Urdu).[2] In Japanese, N+V CPs based on the light verb *suru* are well-documented (GRIMSHAW & MESTER 1988). Tamil, and Dravidian in general, also make extensive use of N+V CPs (ANNAMALAI & STEEVER 1998: 124). Turkish too has N+V CPs, though they tend to be treated as "incorporation" (cf. NILSSON 1986), as are those of Polynesian languages (see ROSEN 1989 – more on this below). We should also mention the CPs of some languages of Australia, analysed in detail for Jaminjung (Non-Pama-Nyungan) in SCHULTZE-BERNDT 2000. I will not be going into the Australian data here because the CPs do not consist of noun-plus-verb combinations, but of coverb-plus-verb, a somewhat different category. In the Iranian languages, to which Kurdish belongs, N+V CPs are ubiquitous (for Persian see TELEGDI 1955, MEYER-INGWERSEN 1974, SHEINTUCH 1976 and WINDFUHR 1979: 113–120). However, there has been no discussion of Kurdish CPs in a modern linguistic framework, although they are briefly touched on in traditional grammars (e.g. BEDIR KHAN & LESCOT 1986: 160–165). Obviously this brief summary is not exhaustive. It is merely intended to embed the Kurdish data in a broader typological context.

3. Kurmanjî Kurdish: background, data, structure

3.1. The Kurdish language and use of sources

Kurdish is a Northwest Iranian language spoken throughout Southeastern Turkey, Northern Iraq, and in parts of Iran, Syria and Armenia. There are two main dialects, Soranî (or Central Kurdish) and Kurmanjî (or Northern Kurdish). Mutual intelligibility between the two varies from non-existent to moderate, depending on the amount of exposure speakers have had to the other dialect. This paper is based exclusively on data from Kurmanjî Kurdish; however I will be referring simply to Kurdish, except where it is necessary to make a distinction.

Kurdish is not a national language anywhere, though it has had at various times some status as a minority language in education. Written forms exist, using the Cyrillic alphabet in the states of the ex-Soviet Union, the Arabic alphabet in North Iraq and Iran, and the Roman alphabet in Turkey. Among the Roman-based publications, a broad consensus now exists on many issues of standardization (see MATRAS 1989). It is important to note that the overwhelming majority of Kurds receive their formal education in childhood through the medium of another language, while literacy in Kurdish is a skill learned in adulthood, and usually in the diaspora. Thus while Kurmanjî Kurdish is the natural L1 of many millions of speakers, it is the "written L1" for very few. Today's emergent standard is largely determined by speakers who in a very real sense are learning as they go (DORLEIJN & LEEZENBERG 2000). As a result, judgements on certain constructions often vary dramatically from speaker to speaker. This is not of course unique to Kurdish–English speakers vary considerably in their judgements too. But there is nevertheless a difference in degree between the status and use

[2] VV sequences with aspectual meanings are not found in Kurdish, although the periphrastic causative construction with *dan* 'give' is structurally very similar to the Urdu "permissive construction", treated by BUTT (1997) as a complex predicate.

of a language like English, the product of centuries of standardization and the basis of mass education, and that of Kurdish, a recently emergent standard lacking a political centre. An additional problem in evaluating Kurdish data is that the repressive language policies in Turkey, the traditional homeland of most Kurds, have made language a highly sensitive political issue.[3] Informants may reject perfectly natural utterances not on grammatical grounds but because they feel they are not "pure Kurdish".

For the above reasons, I have opted for a data-driven analysis of CPs: the primary data stems from a corpus of contemporary literary and journalistic Kurdish (see Appendix 1). For some supplementary issues, acceptability judgements were obtained from two native speakers, but I have taken care not let crucial aspects of the analysis hinge on such judgements.[4] Despite some variation in the data, within the realm of complex predicates certain invariable structures can be discerned, and it is with these that I will be concerned.

3.2. Transitivity and alignment in Kurdish

The notion of transitivity is crucial to the Kurdish verb system and I will therefore begin by clarifying my use of the term. In Kurdish, it is necessary to distinguish at least two different types of transitivity. The reasons for this are best illustrated using examples of CPs:

(2) *wî* *jî* *can* *da [...]* [GIR: 15]
3S:OBL too soul give:PST(3S)
'he too died' [...] (lit. 'he too gave soul')

(3) *Bihar-ê* *dest* *pê* *kir-i-ye* [UZUN: 9]
spring-OBL hand to.it do:PST-PF-3S
'Spring has begun' (lit. 'Spring has put hand on it')

The CP in (2) is *can dan* 'die', literally 'soul give', and in (3) it is *dest pê kirin*, 'begin', literally 'hand to-it put'.[5] Semantically, both clauses express single-participant events involving no control, volition, punctuality, and no transfer of activity, i.e. events as semantically intransitive as one could imagine (cf. HOPPER & THOMPSON 1980). Nevertheless, the verbs *kirin* and *dan* are lexically transitive, and trigger the ergative construction (cf. the oblique case on the subjects). To what extent the clauses (2) and (3) as a whole can be considered transitive is an issue we will tackle below. But at any rate we can note that although these clauses display ergative alignment, generally associated with transitivity, semantically they express intransitive propositions.

Mismatches of this kind make it necessary to draw a terminological distinction between lexical transitivity, a property of individual simplex verb lexemes, and clausal or syntactic transitivity, a property of clauses.[6] Lexical transitivity is an invariant feature of individual

[3] For details on Turkey's policy on Kurdish see SKUTNABB-KANGAS & BUCAK 1995 and HAIG 2001.

[4] I am extremely grateful to ISMET RAMM and ISMAIL SEVER for their unfailing patience and good-humour during many hours of discussions on Kurdish.

[5] The word *pê* is a dummy prepositional object (a contraction of the preposition *bi* and the oblique pronoun *wê/wî*). CPs with dummy prepositional objects are ignored henceforth because they involve very specific problems beyond the scope of this paper.

[6] Multi-level approaches to transitivity have been utilized in the analysis of several other languages, e.g. Saliba (Oceanic, Papua New Guinea, discussed in MARGETTS 1999: 39–49), Ingush (North Central Caucasian, Caucasus, (NICHOLS 1975), or Zazaki (Anatolia, Iranian, PAUL 1998).

verb lexemes, hence comparable to, say, the feature of gender on nouns. It is manifested in the formal features associated with a particular class of verbs. In Kurdish, the crucial defining feature of transitive verbs is that they have ergative alignment in past tenses, but accusative alignment in present tenses. In other words, the alignment of case and agreement patterns with transitive verbs is tense-sensitive. Intransitive verbs on the other hand have a single uniform alignment regardless of tense (see MATRAS 1992/1993, MATRAS 1997, DORLEIJN 1996 and HAIG 1998 for outlines of ergativity in Kurmanjî Kurdish).

The canonical ergative construction in Kurmanjî Kurdish is characterized by three features. I adopt the conventions of DIXON 1994 in referring to the transitive subject as A, the transitive object as O, and the intransitive subject as S:

(4) Formal features of the canonical ergative construction
  (i) Person agreement on the verb: The verb agrees in number and person with the transitive object (O)
  (ii) Case marking of the O: The O is in the absolute case.
  (iii) Case marking of the A: The A is in the oblique case.

These properties are illustrated in (5): the oblique case of the A (*wan*), the absolute case of the O (*ez*), and the agreement of the verb in person and number with the O (1s):

(5) *Wan* *jî* *ez* *dît-im* [MEY: 21]
3PL:OBL too 1S see:PST-1S
'They too saw me'

The point that needs emphasis in the present connection is that the ergative construction is solely determined by the lexical specification of the governing predicate *vis-à-vis* the feature [±transitivity]; semantic characteristics of the clause, such as volitionality or control (which play a role in e.g. Hindi ergativity) are irrelevant. A verb that is lexically specified as transitive triggers ergative alignment, regardless of context.

Now of course transitive verbs show various other semantic and behavioural properties, typical of transitive verbs across languages. For example, many of them refer to two-participant states of affairs; some of them imply control on the part of the A; many of them govern an O; and most of them cannot take the transitivizing suffix *-and*. However, none of these properties is sufficient to account for all members of the class, and some of these properties also apply to verbs that are lexically intransitive. Therefore, the more rigorous approach to defining transitivity is in terms of the **tense-sensitivity of alignment**. Accordingly, we define transitive verbs as follows:

(6) Definition of transitive verb
With transitive verbs, the case-marking of core arguments is sensitive to tense alternations on the verb. With intransitives it is not.[7]

[7] Readers may wonder why the definition does not simply state that transitive verbs trigger ergative alignment. The reason for the present wording is that texts contain many examples of past transitive clauses which deviate from the canonical ergative construction outlined in (4) and (5). As DORLEIJN (1996) shows, in some varieties of Kurmanjî the features of nominal case and verbal agreement are to a remarkable extent independent of one another. In the texts I have dealt with, the sole feature which is shared by all ergative constructions is the **oblique case of the A**, hence the wording of the definition in (6).

3.2.1. The interaction of lexical and syntactic transitivity

In contrast to lexical transitivity, syntactic transitivity is a property of clauses. A clause is transitive if it permits the expression of a direct object NP (direct objects may of course be omitted for pragmatic reasons; the definition therefore hinges on whether a direct object **can** be supplied without impairing grammaticality, not on its actual presence). But what is a direct object? Direct objects display the following formal properties in Kurdish:

(7) Properties of prototypical direct objects in Kurmanjî:
- (i) Direct objects are non-adpositional.
- (ii) In pragmatically neutral clauses, direct objects are immediately pre-predicate.
- (iii) With accusative alignment, the direct object takes the oblique case (unless used in a generic, or sortal sense)[8]
- (iv) With ergative alignment, the direct object determines person and number agreement on the predicate (cf. (5))
- (v) Only direct objects achieve subjecthood under passivization (cf. RAMM 2000)

Examples of direct objects with simple verbs are given in (8) and (9):

(8) *Ez ê nuha **cêr** da-gir-im û b-ê-m* [MEY: 64]
1S FUT now pot:OBL PV-take:PRES-1S and IRR-come:PRES-1S
'Now I will (go and) fill the pot and come back'

(9) *Ez ê **hedîye-k-î** bi-d-im te [...]* [MUŞ]
1S FUT gift-IND-OBL IRR-give:PRES-1S 2S:OBL
'I will give you a gift [...]'

Indirect objects on the other hand are either adpositional, or are immediately post-predicate and oblique, as in *te* 'you' in (9).

Our definition of syntactic transitivity hinges on the potential occurrence of a direct object in the clause, whereas the class of lexical items identified as "transitive verbs" are defined in (6) in terms of the sensitivity of case marking to tense alternations. How does lexical transitivity interact with syntactic transitivity? The relationship is fairly straightforward, and can be summed up as follows:

(10) Only transitive verbs can govern a direct object; intransitive verbs cannot. However, not all transitive verbs govern direct objects.

In other words, while the presence of a direct object presupposes that the verb is lexically transitive, the converse does not hold: lack of a direct object is no guarantee that a verb is intransitive. To illustrate this, consider the following examples:

(11a) *ez bi Tirkî ni-zan-im*
1S to Turkish NEG-know:PRES-1S
'I don't know Turkish'

(12a) *ez bi wan mêr-an na-ken-im*
1S to DEM:PL:OBL man-PL:OBL NEG-laugh:PRES-1S
'I don't laugh at those men'

[8] It is not known under precisely which conditions direct objects are unmarked for case, but it is clearly linked to factors such as definiteness and genericity.

Neither of the verbs *zanîn* 'know' and *kenîn* 'laugh, smile' govern a direct object. Rather, both have a prepositional complement with the preposition *bi* 'to/through'. Both clauses are therefore syntactically intransitive, and appear to have identical argument structures. However, converting the two clauses into the past tense yields an immediate difference:

(11b) *min* *bi tirkî* *ne-di-zanî* [MEY: 6]
1S:OBL to Turkish NEG-PROG-know:PST(3S)
'I didn't know Turkish'

(12b) *ez* *bi* *wan* *mêr-an* *ne-kenî-m*
1S to DEM:PL:OBL man-PL:OBL NEG-know:PST-1S
'I didn't laugh at those men'

Whereas the form of the subject in (12a) and (12b) remains identical, between (11a) and (11b) there is a change in the case of the subject, and in the agreement pattern on the verb. In other words, case and agreement with the verb *zanîn* is tense-sensitive, while with *kenîn* they are not. Although the two verbs appear to govern an identical syntactic structure, they nevertheless belong to two distinct classes. Thus argument structure is not always a reliable indicator of transitivity[9].

To briefly recapitulate, we have established two things in this section: First, while lexical and syntactic transitivity are clearly related, there is not a one-to-one matching – cf. (11) and (12). Second, a verb that is specified as lexically transitive retains its transitivity feature, regardless of the semantics of the clause in which it is instantiated – cf. (2) and (3).

4. N+V CPs in Kurdish

Kurdish contains a number of conventionalized predicative expressions consisting of a single verbal lexeme, plus one or more non-verbal lexeme. We will henceforth refer to them as CPs. Kurdish CPs can be classified on the basis of two parameters: The lexical class of the non-verbal element, and the transitivity of the verbal element. Turning to the lexical class of the non-verbal element, there are three main possibilities:

1. The non-verbal element is recognizably a noun, as in *derman kirin* 'treat, heal', lit. 'medicine/remedy do/put'.
2. The non-verbal element is recognizably an adjective, as in *pîroz kirin* 'celebrate', lit. 'celebratory do/put'.
3. The non-verbal element cannot be identified with a noun or an adjective in the language. An example is the verb *dagirtin* 'fill' in (8), from the particle *da* and the verb *girtin* 'take, grasp'. The non-verbal elements have very heterogeneous sources; some are etymologically related to adpositions, some to adverbs, while some display no discernible relationship to another lexeme of the language. This class has traditionally been termed "particle" or "preverbs" (glossed PV).

[9] Occasionally *zanîn* 'know' does take a direct object, with no perceptible change in meaning. However, the construction with a prepositional object is far more usual.

The second parameter, that of the transitivity of the verbal element, involves a two-way distinction between transitive and intransitive, justified in the preceding section. Combining these two parameters yields a six-way classification, given in Table 1:

	NON-VERBAL ELEMENT IS:		
VERB IS:	PARTICLE	ADJECTIVE	NOUN
INTRANS.	*derketin* 'go out' (particle *der* + verb *ketin* 'fall')	*rast hatin* 'encounter' (adjective *rast* + verb *hatin* 'come')	*dilketin* 'fall in love' (noun *dil* 'heart' + verb *ketin* 'fall')
TRANS.	*vekirin* 'open' (particle *ve* + verb *kirin* 'do, put')	*pîroz kirin* 'celebrate' (*pîroz* 'celebratory' + verb *kirin*)	*maç kirin* 'kiss', cf. (1)

Table 1: Taxonomy of different types of complex predicates[10]

In the interests of brevity, this paper deals only with CPs consisting of a noun and a transitive verb, i.e. the bottom right-hand cell in Table 1.

N+V CPs in Kurdish constitute a large and growing class, with individual CPs differing from one another in terms of degree of idiomaticization. Often the N is a borrowing, but this is by no means always the case. Like other West Iranian languages, Kurdish lacks morphological means for deriving new verbs from nouns (cf. Windfuhr 1979: 113 on Persian), so CP-formation is the sole means of creating new verbal expressions. The following list of 25 CPs cannot therefore claim to be anything like exhaustive. They were chosen because they are all well-established, i.e. familiar to the majority of speakers, attested across most dialects and listed in most dictionaries, and represent the most important structural types:

(13) Examples of N+transitiveV complex predicates

COMPLEX PREDICATE	N	V	GLOSS
a. *alîkarî kirin*	*alîkarî* 'help'	*kirin* 'do/put'	help
b. *av dan*	*av* 'water'	*dan* 'give'	irrigate/water
c. *bang kirin*	*bang* 'call, voice'	*kirin* 'do/put'	call
d. *bar kirin*	*bar* 'burden, load'	*kirin* 'do/put'	(a) load (b) *mala xwe* ~ move (house)
e. *bawer kirin*	*bawer* 'belief'	*kirin* 'do/put'	believe
f. *baz dan*	*baz* 'jump, running'	*dan* 'give'	(a) run away (b) jump (over)

[10] Table 1 ignores a number of complications, for example (i) CPs which contain more than one non-verbal element, as in the CP *dest pê kirin* 'begin' in (3); (ii) CPs where the non-verbal element occurs after the predicate, as in *anîn cîh* 'fulfill', lit. 'bring to (its) place'; and (iii) difficulties in distinguishing adjectives from nouns – see § 6.

COMPLEX PREDICATE	N	V	GLOSS
g. *behs kirin*	*behs* 'subject, topic'	*kirin* 'do/put'	talk about, mention
h. *çap kirin*	*çap* 'print'	*kirin* 'do/put'	publish
i. *deng kirin*	*deng* 'voice, vote'	*kirin* 'do/put'	call
j. *derman kirin*	*derman* 'medicine, treatment'	*kirin* 'do/put'	treat medically
k. *dest avêtin*	*dest* 'hand'	*avêtin* 'throw'	reach for, grasp, tackle
l. *fahm/fêm kirin*	*fêm* 'understanding'	*kirin* 'do/put'	understand
m. *gazî kirin*	*gazî* 'call, shout'	*kirin* 'do/put'	call
n. *guhdar(î) kirin*	*guhdar* 'listener'	*kirin* 'do/put'	listen
o. *hez kirin*	*hez* 'love, pleasure'	*kirin* 'do/put'	like, be fond of
p. *maç kirin/ paçî kirin*	*maç* 'kiss'	*kirin* 'do/put'	kiss
q. *mêze kirin*	*mêze* 'look, looking'	*kirin* 'do/put'	watch, look at
r. *nişan dan*	*nîşan* 'sign, mark'	*dan* 'give'	show
s. *qal kirin*	*qal* 'subject, theme'	*kirin* 'do/put'	talk
t. *qebûl kirin*	*qebûl* 'acceptance'	*kirin* 'do/put'	accept
u. *şerm kirin*	*şerm* 'shame, shyness'	*kirin* 'do/put'	be ashamed
v. *telefon kirin*	*telefon* 'telephone'	*kirin* 'do/put'	phone
w. *xwey dan*	*xwey* 'sweat'	*dan* 'give'	sweat
x. *xwedî kirin*	*xwedî* 'owner'	*kirin* 'do/put'	bring up, raise
y. *zor birin*	*zor* 'force, violence'	*birin* 'take'	defeat

By far the commonest transitive verb in Kurdish CPs is *kirin* 'do, put'. Indeed, the whole issue of CPs is heavily restricted to a small set of lexical verbs, as it is in other languages (e.g. Hindi/Urdu, Japanese, and Turkish).

Turning briefly to the semantics of CPs, in most cases the meaning is largely predictable from the semantics of the N. The semantic contribution of the V on the other hand is vague; there is no implication of volitionality or control (cf. (13) u. and w.). Thus the V functions largely as a verbalizer, a "light verb" in the terminology of GRIMSHAW & MESTER 1988. However, the verb *kirin* is not restricted solely to forming CPs; it also occurs as a full verb in the sense of 'do, carry out':

(14) *Mirov-ê* *ku* *vî* *kar-î* *bi-k-e [...]* [MEY: 20]
person-L:M COMP DEM:OBL work-OBL IRR-do:PRES-3S
'The person who would do this work [...]'

Although the CPs listed in (13) appear to be superficially similar, closer examination reveals that they belong to two distinct structural types, as well as several subtypes. The feature distinguishing these different types is the syntactic status of the N, and I shall therefore begin the analysis with this aspect before going on to present the different types of CP.

At first glance, the Ns appear to be good candidates for direct object status. They are non-adpositional, and occur immediately to the left of the V. However, on the other criteria mentioned in (7), they fail to varying degrees. First of all, they cannot take the oblique case, or if they do, the clause undergoes a significant semantic change. Compare (15a) and (15b).

(15a) *bawer ke!* 'believe!' (= imperative of (13e))

(15b) **vî* *bawer-î* *ke!* intended meaning: 'do/hold this belief!'
DEM:OBL belief-OBL

A further difference between the N of the CP and a prototypical direct object is that the N cannot take the indefinite suffix *-ek*. Compare (16a) and (16b):

(16a) *Wî* *pir* *ji* *Bozo* *hez* *di-kir.* [GIR: 45]
3S:OBL much from Bozo love PROG-do:PST(3S)
'He loved Bozo very much.'

(16b) **Wî* *ji* *Bozo* *hez-ek* *di-kir.*
3S:OBL from Bozo love-IND PROG-do:PST(3S)
Intended meaning: 'He did a little loving of Bozo.'

Similarly, the Ns in these CPs do not generally permit expansion with nominal modifiers (e.g. adjectives), or with quantifiers.[11] Compare the following two examples:

(17) *Ez* *ê* *birîn-a* *te* ***derman*** *bi-k-im*
1S FUT wound-L:F 2S:OBL medicine IRR-do:PRES-1S
'I will treat your wound'

(18) **Ez* *ê* *birîn-a* *te* ***derman-ek-e*** ***baş*** *bi-k-im*
1S FUT wound-L:F 2S:OBL medicine-IND-L:F good IRR-do:PRES-1S
intended meaning: 'I will do a good treatment to your wound'

Finally, the evidence from passivization strongly suggests that these Ns are not true direct objects. The passive in Kurmanjî Kurdish is fairly straightforward. It is based on a form of the verb *hatin* 'come' combined with the infinitive of the main verb. A direct object, and only a direct object, advances to subjecthood, and the original agent is deleted (agented passives have been introduced into the emergent language of the press; they are virtually non-existent in natural spoken discourse). Three examples of passivized CPs should suffice (in the interests of brevity I will dispense with a full glossing):

(19) Passive from the CP *av dan* 'irrigate' (lit. 'water give'):
Bi xwînê û rondikên çavan ***hatine avdan*** *û avakirin.* [AW: 72]
'(they) **have been irrigated** and built up with the blood and tears of (their) eyes'

(20) Passive from the CP *hez kirin* 'like, love'
Bi wî awayî, lawê Zoro Axa wisan di gund de ***hate*** *naskirin û* ***hezkirin***[12] [GIR: 38]
'In this manner the son of Zoro Axa **had become** so well-known and **loved** in the village [...]'

[11] One group of CPs, however, regularly permits genitive attributes of the N – see § 4.3.

[12] In fact the passive of *hezkirin* 'like, love' is unusual, because this CP requires a prepositional complement rather than a direct object. I suspect that it's occurrence in this example is linked to the fact that it is coordinated with another passivized CP, *naskirin*, which does take a direct object. Nevertheless, (21) remains problematic.

(21) Passive from the CP *derman kirin* 'treat'
Cahsên[13] *birîndar li nexweşxaneya leşkerî ya Colemergê* ***têne dermankirin.*** [CTV: 41]
'The wounded persons **are being treated** in the military hospital in Colemerg'

The subjects of these passive sentences is not the N of the CP. In (19), the subject is a discourse recoverable 'they' (= towns and villages from the preceding context), in (20) it is Zoro Axa's son, and in (21) it is the wounded persons. Note that in these examples the N is written together with the V, i.e. the CP forms a single orthographic word. We can therefore safely conclude that the Ns of these CPs do not become grammatical subjects under passivization, but remain amalgamated with the V.

The evidence from case-marking, compatibility with the indefinite suffix, expansion with modifiers, and behaviour under passivization indicate that the Ns of certain CPs lack some of the criterial properties of direct objects. It should, however, be noted that many finer-grained differences exist in the syntactic properties of the Ns of different CPs, which for reasons of space cannot be addressed fully here. We can nevertheless conclude from the initial evidence that the Ns in CPs differ significantly from fully fledged direct objects.

4.1. The fundamental split: Incorporating versus non-incorporating CPs

The evidence from the tests in the previous section suggest that the N of the CP lacks some of the properties of direct objects. However, CPs differ significantly among themselves in just how much objecthood the N possesses. The clearest diagnostic for degree of objecthood utilizes a very general constraint on Kurdish clauses, given in (22):

(22) Each clause can contain only one exponent of the direct object category

(22) is essentially a manifestation of the theta-criterion in GB. Similar constraints on direct objects obtain in many languages, e.g. Japanese (SHIBATANI 1973: 341–342), (BAKER 1988: 177). Now if the Ns in the CPs listed under (13) are direct objects, then the clauses should not be able to accommodate an additional direct object NP, because this would violate (22). But as we have seen, several CPs do permit additional NPs. For example in (1), repeated here for convenience, the direct object is *destê melê* 'the mollah's hand':

(23) *Ferho* ***dest-ê*** ***mel-ê*** *maç* *kir* *[...]* [= ex. (1)]
Ferho hand-L.M mollah-OBL kiss do:PST(3S)
'Ferho kissed the hand of the mollah' (lit. 'kiss-did the hand of ...')

Likewise, example (17) has a direct object *birîna te* 'your wound'.

The facts from verbal agreement lend further support to the conclusion that such NPs really are direct objects. Recall that in the ergative construction, person agreement on the verb is controlled by the direct object. This was illustrated with (5), which I repeat here:

(24) *Wan* *jî* ***ez*** ***dît-im*** [= ex. (5)]
3PL:OBL too 1S see:PST-1S
'They too saw me'

[13] I assume that this word is a rendering of Turkish *şahıs* 'person' (itself a borrowing from Arabic).

The direct object *ez* 'I' determines the person agreement on the predicate. Consider now (25) with the CP *qebûl kirin* 'accept'. Here agreement on the verb is not determined by the N *qebûl*, but by the additional object *ez* 'I':

(25) *Ew-î* ***ez*** *gelek bi xweşî qebûl* ***kir-im*** [...] [ŞK: 46]
3S-OBL 1S much with pleasure acceptance DO:PST-1S
'He accepted/welcomed me (into his home) with great pleasure [...]'

Thus at least some CPs permit the expression of a genuine additional direct object. Assuming that the constraint against double objects given in (22) is correct, if a CP permits an additional direct object then its N cannot also be a direct object. In terms of phrase structure, it appears to be incorporated into the predicate.

However, not all CPs behave in this manner. A large group do not permit an additional direct object. Examples are *hez kirin* 'like, love', or *mêze kirin* 'look at':

(26) *Avukat [...] li saet-ê mêze kir* [GIR: 85]
Lawyer at clock-OBL looking DO:PST(3S)
'The lawyer [...] looked at the clock'

CPs such as *hez kirin* 'love' and *mêze kirin* 'look' can license additional arguments, but crucially, they are not direct objects, but adpositional objects (cf. the preposition *li* 'at' in (26)). Unlike direct objects, prepositional objects never control verbal agreement. In fact, CPs such as *mêze kirin* are ungrammatical with additional direct objects.

Thus on the basis of the criterion "ability to govern an additional direct object" we are able to identify two groups of CPs. For reasons that will become apparent, I will refer to these two types as **incorporating CPs** (permit an additional direct object) and **non-incorporating CPs** (do not permit an additional direct object). Among the CPs listed under (13), the following are incorporating CPs:

(27) Examples of incorporating CPs from (13)

derman kirin 'treat'	*qebûl k.* 'accept'	*nişan dan*[14] 'show'
çap kirin 'publish'	*av dan* 'irrigate'	*xwedî k.* 'bring up'
maç kirin 'kiss'	*fehm k.* 'understand'	*bar k.* 'move, carry'

From a semantic point of view it is not readily apparent why CPs such as *derman kirin* 'treat' and *qebûl kirin* 'accept' take additional direct objects, while *hez kirin* 'love' and *mêze kirin* 'look' cannot. The variation appears at present to be unpredictable, a point made in the same connection for Hindi, which shows a similar split, by MOHANAN (1997: 466). We will therefore simply accept for the time being the existence of two structurally distinct types of CP, and proceed to analyse each in turn.

[14] The syntax of *nişan dan* 'show' is rather more complex than the others. It is actually a hybrid in that it allows expansion of the N *nişan* with a possessor, but nevertheless permits an additional direct object. For the latter reason I include it in this group.

4.2. Incorporating CPs from a typological perspective

We have established that the N in incorporating CPs does not have argument status in the clause; it is syntactically inert, hence the verb is free to assign the direct object role to another NP. Notice that my use of the term incorporation here relates to loss of syntactic status. One could of course take issue with this usage. GERDTS 1998, for example, states that true incorporation involves a noun stem, not a noun, and presupposes morphological incorporation into the host verb. As will become clear in § 4.4, in Kurdish the noun is not morphologically part of the verb. Other labels would therefore be equally appropriate, e.g. noun stripping, or even compounding. Nevertheless, I will continue to use the term incorporation, because it has been applied to similar processes in, for example, German (cf. GALLMANN 1999). But I would like to emphasise that my use of the term "incorporation" really means "syntactic incorporation", i.e. the loss of all argument status.

A formal representation of this type of incorporation is given below: In (28), the structure of a more simple transitive clause headed by the verb *kirin* in the sense of 'do, carry out' is given; (29) represents the structure of an incorporating CP with *kirin* (V represents the lexical verb, the square brackets enclose the predicate complex, the pointed brackets enclose the core arguments):

(28) $<x, y>\ [V]_{tr}$ *kirin* in simple transitive clause (cf. 14)

(29) $<x, z>\ [yV]_{tr}$ *kirin* in an incorporating CP (cf. 1)

The crucial point about this type of incorporation is that the verbal complex in (29) retains its lexical transitivity (i.e. continues to trigger ergativity), and more importantly, the entire verbal complex in (29) displays syntactic transitivity, i.e. is capable of governing an additional direct object, represented in (29) as "z".

This concept of incorporation is radically different to that found in many languages. Elsewhere, we find a noun is incorporated into a transitive verb, and the result is an intransitive verb. An example from Samoan (Polynesian, Oceanic) from MOSEL & HOVDHAUGEN (1992: 738) illustrates this better-known type of incorporation (see also BAKER 1988: 88 for further examples and discussion):

(30) *Na* *fa'atau* *e* *le* *tama* *le* *pua'a*
PST sell ERG ART boy ART pig
'The boy sold the pig'

(31) *Fia* *faatau* *puaa* *oe?*[15]
want sell pig 2s
'Do you want to sell pigs?'

In (31), the noun *puaa* 'pig(s)' has a generic sense. It lacks an article and has been incorporated into the predicate complex. The resultant clause is intransitive, clearly shown by the fact that the subject *oe* 'you' does not have the ergative particle *e*, present in the transitive clause given in (30).

[15] These two examples reflect different orthographic conventions in Samoan, hence the minor discrepancies.

This type of incorporation can informally be represented as follows (the linear order of the constituents is of no concern here):

(32) $<x, y> [V]_{tr}$ Non-incorporated structure (cf. 30)

(33) $<x> [yV]_{intr}$ Incorporated structure (cf. 31)

Note that in contrast to the Kurdish structure given in (29), in (33) the verbal complex is intransitive. GALLMANN (1999: 284) refers to this type of incorporation as "intransitivizing incorporation", ROSEN 1989 talks of "compound noun incorporation". Regardless of the labels, this type of incorporation is completely lacking in Kurdish.

A closer parallel to the Kurdish type is what ROSEN 1989 terms "classifier noun incorporation", widespread in Native American languages and some languages of Australia. In this construction, a transitive verb incorporates a noun, but remains transitive (ROSEN 1989: 302). Furthermore, the resulting complex may govern an additional direct object, what Rosen refers to as "doubling". Clearly this is close to what we find in Kurdish. However, there is a crucial difference: doubling is only possible when the incorporated N is semantically linked to the external direct object. For example, a possible construction of this type could be rendered 'he animal-picked up a dog', where the external object *dog* belongs to "the class of items specified by the incorporated noun", *animal* (ROSEN 1989: 303). Another possibility is that the external direct object is a "stranded" quantifier or determiner of the incorporated N (e.g. 'I men-saw **two**' = 'I saw two men').

The similarities between Kurdish incorporation and "classifier compounding" are limited to the preservation of verb transitivity; the semantic restrictions on incorporation do not hold in Kurdish, as a glance at any of the examples above demonstrates. Nor does Kurdish have any parallels to incorporation involving the stranding of determiners. In fact, I am unaware of any detailed treatment of Kurdish-type incorporation in the literature. Nevertheless, this construction is well-known in the Iranian languages: For Zazaki, PAUL (1998: 131–132) draws an identical distinction between incorporating structures, termed by him "*zusammengesetzte Verben*" and other CPs. In Hindi (Indo-Aryan), a virtually identical structure is found, which both BURTON-PAGE 1957 and MASICA (1991: 368) are careful to distinguish from other types of CP. Likewise MOHANAN 1997 draws a clear dividing line between incorporating CPs and others, although her terminology is different. The phenomenon is thus well-known, but has been treated under a variety of different labels and has not to my knowledge been discussed in the cross-linguistic investigation of incorporation.

Finally, we may ask the question of where and how incorporation in this sense occurs. BAKER's 1988 analysis of incorporation as a syntactic process of X-bar movement, by which the direct object moves into the verb and leaves a trace behind, has been very influential and provides a theoretically attractive solution to the stranding of determiners or quantifiers briefly discussed above. However, the whole debate on whether incorporation is syntactic or lexical presupposes that one should draw a categorical distinction in the first place. Even if one's theory does (e.g. BAKER's version of GB), this is not in itself proof that human language works that way. Other theories, such as construction grammar, approach the issue somewhat differently (cf. SCHULTZE-BERNDT 2000: 23–26 for references). Clearly this is an issue that cannot be resolved here (see ROSEN 1989 for discussion of syntactical and lexical approaches to incorporation). But it should be noted that incorporating CPs in Kurdish clearly have a lexical dimension: they are limited to occurring with just two lexical verbs (*kirin* 'do' and *dan* 'give'), so from the point of view of economy, we would do well to begin by stipulating the peculiar properties of these two verbs in their lexical entries, thus dis-

pensing with much of the proposed syntactic apparatus. The same holds true for comparable constructions in other languages, which are often restricted to just a handful of lexical verbs.

4.3. Non-incorporating CPs

In many respects, the syntax of CPs can be considered as differing solutions to the problem of how to accommodate an additional noun into a simple clause without violating the constraint on double objects given in (22). As we have seen, the N in some CPs is simply ignored; it is syntactically inert, and the clause takes an additional direct object. With non-incorporating CPs, however, no such additional direct object is possible. Basically, we can distinguish three types on non-incorporated CP: (i) saturated CPs; (ii) adpositional complement CPs; (iii) possessor complement CPs.

CPs of the first type require no additional argument. Examples are *baz dan* 'run (away), jump', *şerm kirin* 'be embarrassed', *xwey dan* 'sweat', or *can dan* 'die' (cf. (2)). I will term CPs of this type **saturated CPs**, because the N satisfies the argument requirements of the verb. Saturated CPs express intransitive states of affairs involving a single participant. I will not be concerned further with saturated CPs in this paper, though it should be noted that the difference between them and other CPs is primarily one of the composite semantics of the entire expression. Structurally they behave in a similar manner to other CPs (e.g. the verb remains lexically transitive and triggers the ergative construction in the past tense).

The second type of CP requires an additional argument, that is, is there is a fairly fixed type of oblique complement that they are conventionally associated with. The additional argument is coded as an adpositional phrase. Accordingly, we call this type **adpositional complement CPs**. Typical examples are *mêze kirin* 'look', which requires a prepositional complement with the preposition *li* 'at' (cf. (26)), and *hez kirin* 'like, love, be fond of', which requires a prepositional complement with *ji* 'from':

(34) *Em pir **ji** **te** hez di-k-in* [AW: 71]
1PL much from 2S:OBL love REAL-do:PRES-PL
'We are very fond of you'

A possible analysis of this structure would be to consider the prepositional phrase as a sub-constituent of the N, yielding a PP-complement [*ji te hez*] 'love of you'. On this analysis, we have something like 'do [love of you]'. But an analysis along these lines would be misguided. First, Kurdish does not link prepositional phrases to nouns in this manner; rather, it makes (very occasional) use of post-head prepositional phrases, linked to the head noun via the linking morpheme. Second, the prepositional phrase can be separated from the N via the particle *jî* 'and':

(35) *Ez [...] ji xwe **jî** hez nakim* [AW: 75]
1S [...] from REFL and love NEG-do:PRES-1S
'And I don't love myself'

Finally, the N may actually be preposed to the prepositional phrase, as in (41) below.

We must therefore assume that the prepositional phrase is a genuine argument of the complex predicate, with the same status as prepositional arguments of simple predicates (for

example *pirsîn* 'ask', which takes a prepositional phrase as argument). Clear support for the constituent status of the prepositional complement comes from coordination:

(36) *Sûc-ê* *wan* *ew* *bû* *ku* *wan,* *ji* ***welat-ê***
crime-L:M 3PL:OBL DEM be:PST(3S) COMP 3PL:OBL from country-L:M
xwe], ***[ji*** ***zimanê*** ***xwe]*** *hez* *di-kir*
REFL from language-L:M REFL love PROG-do:PST(3S)
'Their crime was that they loved [their country], [their language]' [AW: 75]

Here two prepositional complements, both dependent on the CP *hez kirin*, are coordinated in a manner typical for arguments of verbs, but impossible for subconstituents of NPs. Thus the prepositional phrases of this type of CP are arguments licensed by the entire CP, not subconstituents of an NP headed by the N. The correct analysis of this type of CP would therefore be along the lines of $[\text{do love}]_{VP}$ $[\text{prep. X}]_{PP}$

The third type of non-incorporating CP also requires an additional argument. However, in this type the additional argument takes the form of a possessor to the N, connected to it via the linking morpheme. I term this type **possessor complement CPs**. An example is the CP *alîkarî kirin* 'help':

(37) *Wî* *zanî-bû,* *ku* *dê* *pîrik-a* *wî* ***alîkari-ya***
3S:OBL know:PST-PPF COMP FUT grandmother-L:F 3S:OBL help-L:F
wî *bi-k-e*
3S:OBL IRR:do:PRES-3S
'He knew that his grandmother would help him (lit. '... do his help')' [MEY 103]

The additional complement, in this example *wî* 'him', is a genuine subconstituent of a complex noun phrase headed by *alikarî* 'help'. NPs of this type are perfectly well-formed, and occur regularly outside of CPs. In this sense, the possessor complements of CPs such as (37) differ crucially from the adpositional complement of CPs such as (36), which I claim are arguments of the entire CP.

Support for this analysis comes from the evidence from passivization. In (38), the CP is *qal kirin* 'mentioning do'. The N takes a possessor subject, i.e. it is a CP of the same type as *alîkarîya kirin* 'help'. Under passsivization, the subject is the entire NP *qala me* 'mentioning of us', not simply the possessor:

(38) *Lê* *tu* *car-an* ***qal-a*** ***me*** *ne-hat-iye* *kir-in.*
but no time-PL:OBL mentioning-L:F 1PL:OBL NEG-come:PST-PF do:PST-INF
'But at no time were we mentioned (lit. '... was mentioning-of-us done')' [AW: 69]

Under passivization, the entire NP consisting of the N plus its possessor advances to subject status, suggesting that it does indeed form a constituent. The construction here is thus 'do $[\text{help-of-X}]_{NP}$'. Here then the additional argument ("X") has a very different syntactic status to the prepositional complements of CPs such as (36).

Note that in possessor-complement CPs the additional complement intervenes between the N and the V: the linear structure is: N(linker)NP V. For this type of CP, an incorporation analysis is clearly misguided. The more difficult question is: how are they to be systematically distinguished from occasional combinations of direct objects plus genitive attribute? There is no easy answer to this question; these constructions vary along a cline of idiomaticization, and at present I see little prospect of finding a precise line of demarcation. However, some of the most strongly conventionalized CPs of this type have developed

structural properties that distinguish them from occasional combinations; I take this matter up in § 4.3.1.

4.3.1. Summary of non-incorporating CPs

We have discussed three types of non-incorporating CPs, distinguished by the type of additional arguments they assign. The CPs from the list given in (13) are distributed as follows across the three types:

Saturated CPs (require no additional complement)	Adpositional complement CPs	Possessor complement CPs
şerm kirin 'be embarrassed' *baz dan* 'jump, spring' *xwey dan* 'sweat'	*hez kirin* 'love, like' *mêze kirin* 'look' *guhdarî kirin* 'listen' *qal kirin* 'discuss' *bawer kirin* 'believe'	*deng kirin* 'call' *gazî kirin* 'call' *qal kirin* 'discuss' *behs kirin* 'mention, discuss' *zor birin* 'defeat' *telefon kirin* 'phone'[16] *alîkarî kirin* 'help'

Table 2: Sub-types of non-incorporating CPs

In reality of course things are not as cut and dried as Table 2 would have us believe. There are various intermediate forms as well as cross-dialectal variation in the syntax of non-incorporating CPs which I will briefly address here. First of all, there are variations on the possessor complement type. In the normal possessor-complement type, the additional argument is a possessor linked to the N with the linking morpheme, which is sensitive to the gender of the N (cf. (39)). However, some CPs use an invariant linking vowel, the so-called "generic ezafe" (MacKenzie 1961). For example, the CP *bang kirin* 'call':

(39) ***Bang-î*** *wê* *di-ke.* [AW: 72]
call-GEZ 3S:OBL:F REAL-do:PRES-3S
'(He) is calling her.'

In some texts, even the generic ezafe is omitted:

(40) *Paşê* ***bang*** *di-ya* *xwe* *kir.* [GIR: 27]
afterwards call mother-L:F REFL do:PST(3S)
'Afterwards (he) called his mother.'

Another hybrid type occurs when the N takes an adpositional complement, but unlike example (34), the complement **follows** the N, thus separating it from its V:

(41) *Wî,* ***hez*** *ji* *gel-ê* *xwe* *di-kir.* [AW: 69]
3S:OBL love from people-L:M REFL PROG-do:PST(3S)
'He loved his people.'

[16] There are regional differences with *telefon kirin*. In the dialect of Muş, it takes an adpositional complement (*ji...re*), in most dialects it takes a possessive complement. Such variation is not uncommon, particularly with recently-coined CPs.

Note that here the other order (i.e. ... *ji gelê xwe hez dikir*) is also possible. Thus there is some degree of variation in the syntax of non-incorporating, some of it dialectal, some perhaps reflecting subtle semantic differences which have evaded me. However, it should be stressed that while variation is common across different subtypes of non-incorporating CPs, the fundamental distinction between incorporating and non-incorporating CPs is remarkably stable. Generally, a CP will belong to only one of the two major types.

4.4. The morphological status of the N in incorporating CPs

In the preceding sections we have addressed the issue of the syntactic status of the N in the CP, concluding that in incorporating CPs it has no syntactic status, while in non-incorporating CPs it has at least residual argument status. The question to be tackled in this section is whether there is a loss of morphological autonomy corresponding to the loss of syntactic status, that is, to what extent the N and the V fuse to form a single word.

We have already encountered examples of non-incorporating CPs where the N is separated from the V by other words (cf. (37), (39)–(41) above), so there appears to be little grounds for assuming morphological incorporation here. Therefore, in this section we will concentrate solely on incorporating CPs and the question of whether the N and the V constitute a single morphological item. Note that this issue is orthogonal to the issue of whether the CP is lexically stored, or whether it is syntactically formed. I adopt a broader view of the lexicon that includes phrasal units. Therefore, whether a CP constitutes a single word is primarily a question of morphological structure, not of lexical storage versus syntactic derivation.

The strongest argument against an incorporation analysis is the fact that some morphological material can be inserted between the N and the V. For example, negation of simple verbs in the present tense is via the prefix *na-*, as in the present negated form of the verb *dîtin* 'see':

(42) *ez wî* ***na-bînim*** 'I don't see him'

When CPs are negated, the negation prefix must occur immediately to the left of the conjugated form of the V, as in the following examples from the CPs *av dan* 'irrigate' and *derman kirin* 'treat':

(43) *av* ***na-dim*** 'I don't irrigate' (not **na-avdim*)

(44) *derman* ***na-kim*** 'I don't treat' (not **na-dermankim*)

Like the negative marker *na-*, the progressive/realis marker *di-* also intervenes between the N and the V of the CP (i.e. *derman di-kim* 'I am treating').

A further diagnostic is nominalization. Under nominalization, the direct objects of transitive verbs must be expressed as possessors, connected to the verb via the linker:

(45) *kuşt-in-a şêr* 'the killing of the lion' (i.e. the lion is dead)[17]
kill-INF-L:F lion

[17] A reading under which the lion is construed as the killer was rejected by my informants.

But when incorporating CPs are nominalized, the N and the V of the CP do not form a possessive construction such as (45). Instead, they form a compound:

(46) *avdan* 'irrigation' vs. *dayina avê* 'the giving of the water'

(47) *dermankirin* 'treatment' vs. *kirina derman* 'the putting of the medicine (somewhere)'

Furthermore, the nominalized CP can take a possessive complement corresponding to its direct object, as in:

(48) *dermankirina birînê* 'the treating of the wound' (cf. (17))

Nominalization therefore treats the N and V as a single unit (this was already evident in the passive forms in (19–21), where the CP is treated as a single orthographic word).

Another morphological process is participle formation. Verbs in Kurdish can form a participle with a resultative sense , created from the past stem of the verb plus the suffix *-î*, for example *man* 'stay' → *mayî* 'being left over'. CPs can also form the basis for participle formation, as in

(49) *birîna* ***dermankirî*** 'the **treated** wound'

In other words, the morphological bracketing in cases like this is:

(50) [dermankir]-î

Evidence from morphology is therefore conflicting: One the one hand, morphological material may be inserted between the N and the V, which suggests that they do not constitute a single word (Kurdish does not have infixes elsewhere). Yet nominalization and participle formation treat the N+V as a single morphological unit.

Conflicting evidence of this type is familiar from German. The German CP *staubsaugen* 'vacuum clean', literally 'dust suck' is treated as a single syntactic constituent, as in *ich werde das Wohnzimmer staubsaugen* 'I'll vacuum the living-room', but in the participle form the prefix *ge-* can intervene between the N and the V: *staubgesaugt* (there is regional variation here, some speakers do accept ***ge***staubsaugt). Thus syntactic unity is not always matched by morphological integrity.

It is worth pointing out that the morphological characteristics of N+V CPs sketched here are identical to those of adjective+V and particle+V CPs (cf. § 4, Table 1), a point I return to in § 6. Here too, morphological integrity is preserved in nominalization and participle formation, but violated by some prefixes. But it is interesting to note that in the Mardin and Midyat regions of Southeastern Turkey, some particle+V CPs have melded to indivisible morphological units. For example, the following form of the particle+V CP *der-ketin* 'come out' is attested (RITTER 1971: 6):

(51) *di-der-keve* '(it) comes out' (Midyat region, modified transcription)

Here the realis prefix *di-* is outside the particle *der*; standard Kurdish would be *der-di-keve*. Similarly, in the Mardin district the particle *ve/va* in the CP *va-xwarin* 'drink' is regularly fully incorporated into the verb, thus creating a new verb root: *çay* ***di-va-xwim*** 'I drink tea' (from my own field notes) as opposed to regular *çay va-**di**-xwim*. The texts in RITTER (1971: 42) even contain an example where a noun, *bang* 'call', is morphologically fully incorporated into the CP *bang kirin* 'call'. Such examples appear to be regionally limited, but they illustrate an

important point: CP formation is a dynamic process, and different dialects can be at different stages in the extent that they treat CPs as single units, both syntactically and morphologically. If we consider the process of CP formation as one of increasing lexicalization of syntactic particle/N/adj.+V structures, then full morphological incorporation resulting in the creation of a new verb stem would appear to represent the final stage.

4.5. Summary of CPs in Kurdish

Beneath the apparently uniform N+V template, CPs reveal a myriad of structural differences. The main result of the analysis is that two types of CP can be discerned, distinguished according to the syntactic status of the N. With incorporating CPs, the N is syntactically inert and fails to control agreement. The CP as an entire unit is syntactically transitive, licensing an additional direct object. Syntactic incorporation, however, is not matched by full morphological incorporation: the N and the V of the incorporating CPs can be separated by certain prefixes. Non-incorporating CPs on the other hand do not permit an additional direct object. We can sum up our classification so far as follows:

(52)

INCORPORATING CPs	NON-INCORPORATING CPs
N is syntactically inert, additional direct object is possible	N blocks the direct object slot, no additional direct object is possible
CP is syntactically transitive	CP is syntactically intransitive
N and V are not morphologically incorporated	

Non-incorporating CPs are clearly modelled on productive syntactic patterns in the language, and I follow most scholars of Iranian languages in rejecting a clear dividing line between occasional N+V combinations, and conventionalized CPs. They differ from canonical transitive clauses in degree of idiomaticity, in the lack of certain argument properties of the N (e.g. case marking), and through their ability to license additional arguments. Furthermore, some develop unique structural features, as was shown in exs. (39)–(40), confirming that frequency of association may lead to the emergence of novel structures.

From a cross-language viewpoint incorporating CPs are the more interesting. Constructions of this type constitute a significant typological feature of Iranian and Indo-Aryan languages which is completely lacking in English, and to my knowledge most of the western Indo-European languages. The concept of syntactic incorporation developed here accounts for a number of facts: First, the lack of case marking, word-order freedom, quantifiers and modifiers on the N. This follows from the fact that it is not an argument. Second, the fact that it does not achieve subjecthood under passivization. Third, that it is ignored for agreement purposes (cf. (25)), and finally, that the clause permits an additional direct object. All these facts follow quite straightforwardly from the assumption that the N does not have argument status in incorporating CPs.

5. Argument-sharing approaches: MOHANAN (1997)

In Hindi, CPs consisting of a noun and a transitive verb are very similar to those of Kurdish. MOHANAN (1997) has proposed an analysis of such constructions based on the notion of argument sharing, according to which the N and the V jointly determine the argument structure of the clause. While her paper encompasses CPs with both intransitive and transitive verbs, for ease of comparison I restrict myself to a discussion of the latter.

MOHANAN is concerned with developing a syntactic and semantic representation capable of capturing what she considers to be the compositional nature of CPs. Her analysis embodies three related claims: First, the N of the CP is predicative, and contributes to the argument structure of the clause; second, although the N contributes to the argument structure of the CP, it may itself be an argument of the V; third, the CP constitutes a phrasal category, and is therefore "put together post-lexically, not lexically" (433, all references in this section refer to MOHANAN 1997, unless stated otherwise). In what follows, only the first two claims are considered, while the question of whether CPs are lexically stored or syntactically derived will not be pursued. As mentioned, I prefer a broader conception of lexicon, one which permits phrasal units as primitives even if they violate lexical integrity (for arguments in favour of this position for CPs see ACKERMAN & LESOURD 1997).

Let us turn to the evidence for the predicative nature of the N in Hindi CPs. The following examples from MOHANAN 1997[18] contain the verb *kar* 'do' (etymologically related to Kurdish *kirin*):

(53) *raam-ne* *apnaa* *homwark* *kiyaa* [434][19]
Ram-ERG self's homework(M) do(PF:M)
'Ram did his homework'

(54) *raam-ne* *mohan-par* *bʰarosaa* *kiyaa* [434]
Ram-ERG Mohan-LOC trust(M) do(PF:M)
'Ram trusted/relied on Mohan'

MOHANAN states that *kar* 'do' licenses two arguments, "a doer" and "a done thing". In (53), these are *Raam* and *homwark* respectively. In (54), however, an additional argument, *Mohan,* is present which is not licensed by the verb. According to MOHANAN, the "only plausible account for this additional argument is that it is licensed by the noun *bʰarosaa* 'trust'. In other words, *bʰarosaa* contributes to the number of arguments in the clause, and therefore participates in predicatehood." (434–435)

On the basis of these and similar examples MOHANAN concludes that the N itself is predicative in the sense that it may assign arguments. The predicative potential of the N is in fact criterial in her definition of CP: "a complex construction in which **two semantically predicative** elements jointly determine the structure of a single syntactic clause" (432, emphasis added).

Let us now turn to the second claim, namely that although the N is itself predicative, it is simultaneously an argument of the V. The evidence for this comes from the agreement patterns on the verb. In past tenses, transitive verbs in Hindi agree with direct objects in gen-

[18] Translations and transcriptions are from MOHANAN, glosses have been modified.
[19] Refers to the page number in MOHANAN 1997.

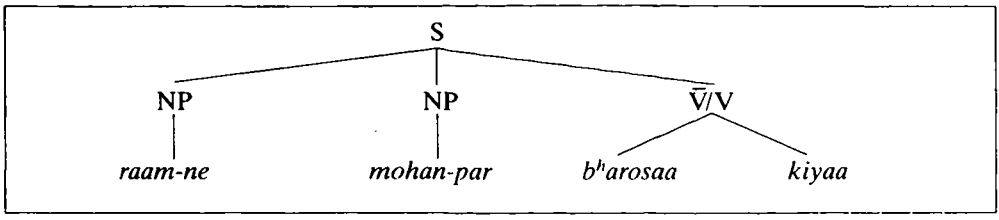
der. For example, in (54), the verb shows masculine agreement. In (55), however, the verb shows agreement with the feminine noun *nafrat*:

(55) *ilaa-ne* *mohan-se* *nafrat* *kii* [454]
Ila-ERG Mohan-INST hate(F) do(PF:F)
'Ila hated Mohan'

The verb thus agrees with the N of the CP in gender. This fact demonstrates that the N must be the direct object of the V, since only direct objects can be unmarked for case, and only direct objects trigger agreement on the verb. Further evidence from passivization (456) apparently supports this analysis. However, the N does not behave in all respects like a fully-fledged direct object. Evidence from scrambling (444–446) reveals that, unlike a normal direct object, the N cannot be separated from the V (except for topicalization, which MOHANAN suggests is distinct from scrambling).

In sum, for CPs of this type we find (a) that the N co-determines argument structure; and (b) that the N possesses at least some of the properties of an argument of V, namely the ability to trigger agreement on the V, and arguably, promotion to subject under passivization. However, it differs from a fully-fledged direct object such as *homwark* in (53) in that it has restricted word order freedom. MOHANAN concludes that the N is therefore "not a direct daughter of S, but part of the verbal constituent" (446). The structure of the CP is represented in the following formalism:

(56) *Raam-ne* *mohan-par* *bʰarosaa* *kiyaa* [= ex. 53]
'Ram trusted/relied on Mohan'



Recall that my analysis of non-incorporating CPs in Kurdish in § 4.1 is in many respects similar: the N lacks the case-marking and freedom of word order normally associated with a direct object, but it nevertheless appears to block the direct object position in the clause, i.e. has some residual argument status (note that for Kurdish, gender agreement on the predicate is not available as a diagnostic of objecthood). In both languages it would appear that an argument may possess some of the properties of an object, but not all, that is, we must assume differing degrees of objecthood.

Let us return to the more central issue, the predicative nature of the N. MOHANAN develops a formalism for representing the predicate structure of the N, the details of which need not concern us here (439–443). Her analysis of the CP *bʰarosaa kar* 'trust' involves the combination of two distinct sets of semantic and argument structure, one for the N and one for the V. These two amalgamate, yielding a monoclausal structure for the CP at the level of argument structure. This is the essence of the argument-sharing analysis.

It is notable that MOHANAN's examples all involve CPs with semantically complex, abstract event nouns such as *bʰarosaa* 'trusting', *safaaii* 'cleaning' or *apmaan* 'insult'. Kurdish too has CPs involving such nouns, for example *bawer kirin* 'believe', *mêze kirin* 'look', or *hez kirin*

'like, love'. The Ns in these CPs can, on semantic grounds, plausibly be considered predicative, and for CPs of this type, the argument sharing approach is certainly an attractive proposal.

5.1. Arguments for not sharing argument-sharing

Although an argument-sharing account offers an insightful analysis of some non-incorporating CPs, it is less clear whether such an approach is viable for all CPs in Kurdish, and indeed in Hindi. There are three main objections. The first concerns MOHANAN's treatment of genitive complements. Some CPs in Hindi have an additional genitive-marked argument, as in the following example:

(57) *raam-ne* ***kamre-kii*** *safaaii* *kii* [437]
Ram-ERG **room-GEN** cleaning(F) do(PF:F)
'Ram cleaned the room'

According to MOHANAN, the genitive marked noun in (57) is an argument of the entire CP, on a par with the locative argument in (54). Thus in the corresponding tree diagram, it is treated as a direct daughter of S (448). Later in her paper (460), however, MOHANAN concedes that genitive nominals are modifiers of the N, and agree with them in gender, hence suggesting that they are not actually direct daughters of the S-node. Despite this, she maintains that the genitive is an independent syntactic argument because it can be scrambled away from its governing noun, and because of its behaviour under passivization.[20] However, another analysis of these constructions, BURTON-PAGE (1957: 477), comes to the opposite conclusion, namely that the genitive complement is a sub-constituent of the N, not an argument of the predicate. I can only conclude that this issue has yet to be satisfactorily resolved.

MOHANAN's analysis of genitives as immediate arguments of the CP is not applicable to Kurdish. Recall that Kurdish too has examples of this type, as in (37) *alîkarî kirin* 'help'. As I noted, the genitive here appears to be a subconstituent of the N. In contrast to Hindi, genitives in Kurdish can under no circumstances be separated from their head noun, and their behaviour under passivization strongly suggests that they form a constituent with the head. In Kurdish, genitive complements are fundamentally different from other types of arguments of CPs, and require a different underlying representation. It is questionable whether argument sharing is applicable in this case.

The second drawback with the argument sharing analysis is that it rests crucially on the assumption that the N is indeed predicative. As mentioned, this assumption is actually built into MOHANAN's definition of CP (432), which presupposes "two semantically predicative elements". On this definition, if an N is not "semantically predicative", then the construction is not a CP. But what does it mean for a noun to be predicative?

[20] The evidence from passivization is not compelling. The sole example MOHANAN provides is the passive sentence *Ilaa-se mohan-kaa apmaan kiyaa gayaa*, translated by her as 'insult was done to Mohan by Ila'. Using evidence from pronoun deletion and control, she concludes that neither *Ilaa-se* nor *mohan-kaa* can be subjects; therefore *apmaan* 'insult' must be the subject. A solution that she does not discuss is that the subject of the passive clause is the complex NP *mohaan-kaa apmaan* 'the insulting of Mohan', yielding a structure something like '[The insulting of Mohan] was done by Ila'.

MOHANAN does not actually give independent evidence for the predicativity of nouns, i.e. evidence outside of their behaviour in CPs. She nevertheless distinguishes between diadic and monadic nouns, i.e. nouns with differing valencies. For example, *bʰarosaa* 'trusting' is considered to be diadic, while *safaaii* 'cleaning' in (57) is monadic. According to MOHANAN, the valency of a noun can be determined partly on the basis of its semantics:

> "trusting involves the one who trusts, and the one who/which is trusted. Cleanness, in contrast, is a state predicate of a location, and does not necessarily involve any other participant. Thus an important basis for deciding valency is meaning." (442)

These arguments are less than convincing. First of all, the word for 'cleanness' is glossed in all examples as 'cleaning', not 'cleanness'. There is a world of difference between 'cleaning' and 'cleanness' in terms of event semantics (one is process, the other result). Secondly, if the word *bʰarosaa* 'trusting' involves two participants, surely the word *puujaa* 'worship' also does, namely a worshipper and a thing/person worshipped. Yet this word is apparently monadic (443–444). Closer inspection of the examples reveals that the nouns considered monadic are those that govern a genitive complement, while those Ns that assign a different case, e.g. locative, are diadic, although this is not stated explicitly by MOHANAN. But as presumably any noun can have a genitive attribute, this does not seem like a particularly reliable diagnostic for determining the valency of nouns.

The most serious objection to argument sharing concerns its application to incorporating CPs. In the Hindi CPs discussed up until now, the N determines agreement on the V (cf. (55) and (57) above), and is hence an argument. However, scholars of Hindi have long since noted the existence of CPs where the N is ignored for agreement purposes, i.e. is not treated as an argument of the V (e.g. BURTON-PAGE 1957). An example is *yaad kar* 'remember':

(58)	*ilaa-ne*	*kissaa*	*yaad*	*kiyaa*	[459]
	Ila-ERG	incident(M)	memory(F)	(PF:M)	
	'Ila remembered/recalled the incident'				

The N *yaad* 'memory' is feminine, but the verb agrees with the masculine *kissaa* 'incident'. The conclusion that MOHANAN draws from these facts is that *kissaa*, not *yaad*, is the object of the clause. This assumption is supported by the facts of case marking, passivization, negative placement and gapping (459–466). Clearly this second type of CP corresponds closely to what I have termed incorporating CPs in Kurdish: the N is syntactically inert. I will therefore continue to use the term incorporating CP, although this is not the terminology used by scholars of Hindi.

Consider now how an argument-sharing approach can account for case assignment in incorporating CPs such as (58). In this type, the additional argument is nominative/accusative (the nom/acc alternation depends on the animacy of the referent). Now according to MOHANAN, this argument is also assigned by the "host" (459), e.g. the N *yaad* 'memory' in (58). However, MOHANAN provides no independent evidence for nouns assigning structural case, so the claim that the N is a case-assigner remains speculative. More to the point, there is a simpler explanation: it is the V that assigns accusative case. As MOHANAN herself notes, the N is not an argument. Therefore the V is free to assign accusative case to an additional NP. This is essentially the incorporation analysis that I have proposed for Kurdish, and I believe it is applicable to Hindi CPs of this type. Note, however, that on this analysis the only argument assigning-element is the V, which obviously contradicts MOHANAN's definition of CP.

Indeed, the assumption that CP formation in Hindi is dependent on a predicative N is generally difficult to accept. Consider the following N+V construction:[21]

(59) *raam-ne* *kal* *bahut* *kaam* *kiyaa*
Ram-ERG very hard work do(PF:M)
'Ram worked very hard'

The expression *kaam kar* 'work', lit. 'do work' is a strongly idiomaticized expression, treated by FAIRBANKS & MISRA (1966: 31) as a "nonverbal compound verb", corresponding to what other scholars have termed CPs. Yet as we can see from the example, it does not require an additional argument beyond its N. One could supply additional facultative complements, as in "at home", or even a genitive complement of *kaam* 'work', as in 'the work of the hotel' (cf. FAIRBANKS & MISRA 1966: 32). But such facultative elements are hardly grounds for claiming that the N is predicative.

On MOHANAN's approach, we must either assume that the N is predicative, or we must exclude such expressions from the set of CPs. This would be a questionable move, given that combinations such as *kaam kar* share many similarities with CPs like *bʰarosaa kar*, and both are considered CPs by traditional grammarians. Ultimately the problem is that MOHANAN's definition of CP presupposes that the N is predicative, but apart from fairly vague semantic criteria there is no readily applicable diagnostic for distinguishing predicative from non-predicative Ns.

5.2. Summary of argument-sharing approaches

The argument-sharing approach is not new. JACKENDOFF 1974 advanced a formal account for English CPs such as *put the blame* which involved combining distinct argument structures for the N and the V. CATTELL (1984: 5) also refers to the predicative nature of nouns in English CPs such as *make a confession*. Such approaches are motivated by the fact that certain nouns, particularly deverbal nouns with complex event semantics such as *blame*, *demonstration*, *confession* etc. possess an argument structure which, at least at an abstract semantic level, is comparable to that of verbs (cf. GRIMSHAW & MESTER 1989: 210 for references). For CPs containing this type of N, argument sharing is a promising avenue. Likewise it is highly relevant for CPs consisting of two or more verbal lexemes, as in Hindi/Urdu (BUTT 1997), and has been widely applied to verb serialization (cf. VEENSTRA 2000: 868). An argument-sharing approach has also been applied to the combination of coverb and verb in Jaminjung (Non-Pama-Nyungan, Northern Australia). Such an analysis appears well justified for Jaminjung, because as SCHULTZE-BERNDT (2000: 135–145) shows, there is independent evidence for the predicative nature of coverbs (e.g. use as secondary predicates, or in imperatives). In some languages, N+V CPs appear to be limited to Ns which display predicative semantics. This is apparently the case for CPs with the Japanese light verb *suru* 'do': GRIMSHAW & MESTER (1989: 210) note that argument merging, the peculiar property of this type of CP, only happens with Ns that refer to "processes", not to results. Here then semantic restrictions on the N are matched by structural distinctions. In general, then, argument sharing makes sense when the predicative nature of both elements can be established on

[21] This example was kindly supplied by RAMAWATAR YADAV (p.c.).

independent grounds, and when the construction under investigation is only possible with such elements.

But this is not the case for Kurdish and other Iranian languages. Consider for example the Kurdish CP *av dan* 'irrigate', lit. 'water give'. An argument-sharing account would have to postulate a predicative structure for the simple mass noun *av* 'water', something not plausible on semantic grounds. In fact, as far as I can ascertain, in Kurdish CP-formation is not limited to any one particular semantic class of nouns.

Argument sharing is thus a promising avenue for certain constructions treated under the label "complex predicate". But in Kurdish, and in other Iranian languages, CPs are formed through a very general strategy of combining non-verbal elements with a verb (see § 6.3). It is certainly the case that among the non-verbal elements we also find nouns with complex event-semantics. But that is not a prerequisite for the structure.

Finally, my rejection of argument sharing for Kurdish should not be construed as dismissing any effect of the N on the argument structure of the CP. It is quite clear that CPs have argument structures distinct from their basic verbs and in this sense they must be considered composite expressions. But the mere fact that the argument structure of the CP is different to that of the simplex verb does not necessarily lead to the conclusion that the N is an argument assigner. It may result at least in part from the fact that the N – in some CPs at least – occupies an argument position of the V, hence causing a shift in the possible argument configuration of the clause. And as I have mentioned, for incorporating CPs, an argument-sharing approach is not plausible.

6. Towards a unified account of CPs in Kurdish

6.1. The categorial status of the non-verbal element (NVE) in the CP

The taxonomy of CPs in Table 1 follows traditional wisdom in distinguishing three different categories of the non-verbal element: nouns, adjectives and particles. Working on the assumption that this is a valid taxonomy we then proceeded to focus on CPs consisting of N+verb, ignoring for example adjective+verb CPs. While the noun/adjective distinction appears reasonably well-founded in Kurdish (nouns have inherent gender, adjectives do not), there are nevertheless complications which I have played down until now. Consider the following two clauses:

(60)	*ez*	*wî*	*qebûl*	*na-k-im*
	1S	3S:OBL	acceptance	NEG-do:PRES-1S
	'I don't accept that'			

(61)	*ez*	*wî*	*pîroz*	*na-k-im*
	1S	3S:OBL	celebratory	NEG-do:PRES-1S
	'I don't celebrate that'			

In (60), *qebûl* is interpreted as the N of a CP, and indeed, if one consults a Kurdish dictionary (e.g. OMAR 1992, RIZGAR 1993) *qebûl* is listed as a noun with feminine gender. Thus we analyse (60) as an example of incorporating CP, with *qebûl* as the syntactically incorporated N. Example (61) looks superficially very similar. Furthermore, the two constructions behave identically on the morphological parameters discussed in § 4.4. However, according to the

dictionaries, *pîroz* is an adjective, and on the logic behind the classification of CPs given in Table 1, we are obliged to interpret (60) and (61) as two distinct structures.

But the dictionary is not telling the whole truth. Examining the actual usage of *qebûl* in a corpus of contemporary written Kurdish (see Appendix) reveals that it is a frequent word, occurring 97 times. However, of these instances 95 are in the CP *qebûl kirin*, while the remaining two are examples of *qebûl* as a predicative noun (one of which is in a piece of poetry, which does not reflect normal usage). In other words, the word *qebûl* practically does not exist outside the combination *qebûl kirin*, a fact that my informants confirmed. I have found no examples anywhere of it as the head of a NP, with modifiers, articles, or case markers. Nor is it used with specific reference. In other words, it lacks all the trappings of a prototypical noun, both in its formal elaboration, and in its discourse function.

Looked at in this light, the categorical difference between noun and adjective, and with it the categorical difference drawn between N+V CPs and N+Adj CPs, begins to shrink. Apart from the dictionary makers' claims, what evidence do we have that *qebûl* is indeed a noun (and the dictionary makers may have been influenced by the fact that Arabic *qabûl*, the origin of the word, is a noun in Arabic)? If, on the other hand, we assume that the N in this CP has lost a good deal of its "nouniness" then we are provided with a ready account of the syntactic properties of the CP: the N is closer to an adjective (used as an adverb): it cannot take articles etc., can be ignored for agreement purposes and, most significantly, it does not occupy an argument slot.

This line of reasoning can be applied to some other incorporating CPs, for example *rast kirin* 'put right', or *nas kirin* 'know, be acquainted with'. *Rast* has two related meanings: as a noun meaning 'the right thing' and an adjective meaning 'right, correct'. Likewise, *nas* can mean either 'acquaintance' or 'familiar'. When combined with a light verb in a CP, it is to my mind impossible to decide whether we are dealing with N+V incorporating CPs, where the N has lost all argument status, or with adjective+V expressions, where the non-verbal element is not lexically capable of argumenthood. Incorporating CPs could therefore be at least in part explained through a process of lexical decategorization, or loss of nounhood of the N.

However, this explanation does not account for all incorporating CPs. For example, the N in *derman kirin* 'treat medically' regularly occurs as the head of an NP with specific reference with case-marking, indefinite suffix etc. Thus the explanation that the N has lost its nounhood cannot be applied across the board. Nevertheless, the affinity between the N of the incorporating CP, and an adjective functioning as an adverb is certainly close, and may provide a clue as to the diachronic origins of incorporating CPs. It is worth noting that this explanation is even more relevant to incorporating CPs in Turkish, where the dividing line in the lexicon between noun and adjective is weak from the start (Braun & Haig 2000). For Hindi too the noun/adjective distinction may not be always clear. For example, Mohanan (1997: 457) treats the combination *pasand kir* 'liking do' as a CP consisting of a noun (*pasand*) and a verb, while in Masica (1991: 369, ex. 259) exactly the same combination is considered to be an adjective plus verb CP. Thus a certain degree of cross-category shift appears to play a role in incorporating CPs, a point I will develop in the next section.

6.2. CPs as a typological feature of Kurdish

In the foregoing discussion we have identified two major types of CP, labelled incorporating and non-incorporating CPs. While structures corresponding to the non-incorporating CPs may be found in many languages, incorporating CPs are a characteristic trait of Kurdish, and indeed Iranian and Indo-Aryan languages generally. This raises the question of whether we are dealing with two fundamentally different constructions. I believe the answer to this question is no; rather, I believe both types of CP should be considered reflexes of a broader language-specific typological trait, which I will briefly sketch in this section.

Recall that in § 4 I distinguished three types of CP, based on the category of the non-verbal element:

Particle+Verb CPs:	e.g. *der-ketin* 'go out'; *ve-xwarin* 'drink'
Adjective+Verb CPs:	e.g. *pîroz kirin* 'celebrate'
Noun+Verb CPs:	e.g. *av dan* 'irrigate' etc.

In § 6.1, I called attention to the fact that the dividing line between nouns and adjectives is in many cases hard to draw. The same applies to the difference between adjective and particle. A good example is the preverb *çê*, which occurs in two CPs, *çêkirin* and *çêbûn* 'make, complete' and 'be made, turn out right'. It was originally an adjective meaning 'good, whole', and this meaning is preserved in the comparative form *çêtir* 'better'. However, the simple adjective has been almost entirely supplanted by other words (*baş*, or *rind* 'good'), and today the word/particle *çê* is almost entirely restricted to occurring in the two CPs mentioned above. Given these facts it is not easy to say whether *çê* is a particle or an adjective. Similar problems of definition exist for so-called particle verbs in Persian (cf. Telegdi 1955: 74–75).

We must therefore accept some degree of overlap exists between the three categories particle, adjective and noun. Furthermore, all three types of CP behave on the morphological criteria mentioned in § 4.4 identically. Finally, none of them affect the lexical transitivity of the verb (this is in sharp contrast to German or Russian preverbs, which regularly affect the transitivity of the verb). In other words, all three types have a good deal in common. Now our three-way classification of CPs rests solely on the lexical class of the non-verbal element, but as we have seen, the differences between lexical categories are to some extent neutralized (e.g. lack of case, articles etc.) in the preverbal position. I believe therefore that a more adequate description of Kurdish CPs would assume a general CP-structure something like (62):

(62) Generalised CP structure for Kurdish:
[XV], where X is an element of some non-verb lexical or phrasal category.

Syntactically, X may be inert: this is the case with particles, adjectives and the Ns of incorporating CPs. Or it may retain argument status (as in non-incorporating CPs). Or it may be a phrasal category, an NP (this is the case with CPs with possessor-complements). But the structural differences between different types of CP do not match up exactly with the differences in lexical category of the non-verbal element. What we find is that some N+V CPs actually behave very much like adjective+V CPs, while adjective+V CPs are structurally virtually identical to particle+V CPs. The commonalities and differences are summed up in the following diagram, where the different types of CP are positioned on a continuum of degree of fusion between the elements:

	Degree of fusion between N and V Minimum ◄——————————————————————► Maximum				
1	N+V Possessor Complement	N+V Adpositional Complement	N+V Incorporating	Adjective+V	Particle+V
2	Noun			Adjective	Particle
3	Argument, NP	Argument, but lacks NP status?	None		
4	affixes, words	affixes, words (?)		affixes/none[22]	
5	yes	yes (?)		no	

Key: 1 = type of CP; 2 = lexical category of non-verbal element (NVE); 3 = syntactic status of NVE; 4 = NVE can be separated from V by (cf. § 4.4); 5 = NVE is capable of independent reference; ? = does not apply to all members of the category

Table 3: Formal properties of different types of CPs in Kurdish

As is evident from Table 3, no single structural factor corresponds to the traditional distinction between nouns, adjectives and participles. In particular, adjectives and particles behave virtually identically on all structural parameters. Thus it is a clearly a small step from particle CPs to adjective+V CPs.

Recognizing that the differences between lexical categories are to some extent neutralized in CPs provides us with a clue to the diachronic development of CPs in Kurdish. CPs with particles are a feature of all the earliest attested stages of Indo-European, and are a significant feature of many of the daughter languages (Telegdi 1955: 173–175, Watkins 1964). We can therefore assume that they are inherited from the earliest stages of Kurdish. However, particle verbs have undergone very different developments in the individual Indo-European languages. In some they continue to be highly productive, especially Germanic and Slavic, where they fulfill (among many other things) grammatical functions as transitivizers and aspect markers. In Persian, on the other hand, although particle verbs are still present in a number of fixed combinations, they are no longer productive (Telegdi 1955). The drop in productivity of particle verbs has been accompanied by the rise of noun-plus-verb CPs, now by far the most productive means of new verb formation in Persian. In Kurdish we can assume a similar development: particle verbs exist, but are no longer a productive means of forming new predicates, whereas N+V combinations are rampant.

We can assume then that the development of CPs in Iranian began with a construction in which verbs combined with particles, and was gradually extended to include other non-verbal elements, adjectives and nouns:

(63) particle+verb → adjective+verb → noun+verb

[22] Cf. dialectal examples of morphological incorporation in § 4.4.

Accompanying this development is a shift of semantic weight away from the verb, now little more than a vague transitivizer (Kurdish *kirin* 'do') to the non-verbal element. In the N+V CPs we have examined here it is the noun which carries the main semantic content. Generally then, CPs in Kurdish can be considered a development of the ancient Indo-European particle-verb construction. What appears to drive the process is a general propensity for verbs to acquire non-verbal "parasites" which over time may fuse may the verb.

What we have not explained is why N+V CPs should have developed in Iranian and Indo-Aryan languages, but not in the Indo-European languages of Europe, where only particle CPs are widespread. The following four structural properties may have contributed to the developments in Indo-Aryan and Iranian:

1. **OV-word order**. This means that the direct object occupies the immediately preverbal position, i.e. shares this position with the particle of the particle verbs.
2. **Lack of morphological means for deriving verbs from nouns**. In Iranian languages generally noun-to-verb derivation is either weak or entirely lacking. Thus CPs become the major source of new verbal expressions.
3. **Differential object marking**, e.g. the zero-marking of direct objects under certain semantic and pragmatic conditions (e.g. definiteness, topicality or animacy, cf. BOSSONG 1985 for Iranian). This means that certain nouns in the direct object position lose one of the features that distinguish them from adjectives, namely case.
4. **Lack of articles**. Nouns are not obligatorily accompanied by overt determiners, again bringing nouns closer to adjectives and paving the way for a loss in argument status.

In Kurdish and Persian, all four properties are present, as they are in Hindi. In Dravidian languages, also known for their N+V CPs, we find the same bundle of features (e.g. Tamil, ANNAMALAI & STEEVER 1998, Kannada, STEEVER 1998 and Telugu, KRISHNAMURTI 1998). However, these features are not necessary and sufficient conditions for the presence of N+V CPs: Turkish has N+V CPs, although the second feature above does not apply to Turkish. But it is notable that no Indo-European language of Europe displays more than one of these properties, and most have none. Obviously this is far from a full explanation for the distribution of CPs. But I do believe that any explanation must be framed in terms of a conspiracy of factors, rather than any single one. In passing we should also note the areal bias in the distribution of N+V CPs.

The important point in the present connection is that the N+V CPs we have been examining can be most revealingly seen within the larger context of X+V complex verbs. On this view, they are a language-specific development of a very old trait of Indo-European. Note that CP-formation in Iranian languages is not dependent on semantically complex nouns. Therefore, it is questionable whether the analysis of English or Japanese CPs, which has focussed precisely on combinations of semantically complex event-nouns with verbs, will be of much relevance for a uniform account of CPs in Kurdish.

7 Conclusions

At the outset of this paper I mentioned two central issues in relation to CPs: one is the status of the N in the construction: is it an argument, is it incorporated into the verb, or does it have some intermediate status? The second point is how does the CP assign its argument structure. The analysis of N+V CPs reveals that neither question has a uniform answer. CPs

in Kurdish display a fundamental split between incorporating CPs, where the N is syntactically inert, and non-incorporating CPs, where the N fills an argument function. For incorporating CPs, the status of the N is relatively straightforward: syntactically it is ignored, but it remains phonetically and morphologically a word in most contexts. For non-incorporating CPs the answer is somewhat more complex, and several sub-types were identified. In some cases, the N has most of the features of a fully-fledged direct object (e.g. advancement to subject under passivization), in others this feature is lacking; there are finer-grained differences here which slipped through the analytical net. But I would conclude that an adequate description of these CPs must include a graded concept of objecthood.

Turning to the second issue, how CPs assign argument structure, again we have conflicting results for different types of CP. For incorporating CPs I suggest that the sole case-assigner is the V. For some non-incorporating CPs (those with adpositional complements) a convincing case for argument-sharing can be made. CP formation in Kurdish, and in Hindi, is not limited to semantically complex event-nouns, so assuming that argument-sharing is central to CP formation generally means either assuming all Ns are predicative, an option that is counter-intuitive and leads to undue complications with incorporating CPs, or it means excluding large numbers of N+V combinations from the domain of CPs. As the criteria for deciding on the predicativity of a noun have not been spelled out in any detail, this latter option also remains problematic.

In sum, neither an incorporation analysis nor an argument-sharing analysis provides an adequate framework for analysing the full range of N+V CPs in Kurdish and related languages. If typology is about developing viable procedures for cross-language comparison without sacrificing the idiosyncratic genius of individual languages, then we will have to look further afield for an adequate analytical framework. I have made some tentative suggestions towards understanding the Kurdish data as a language-specific extension of a more general typological trait, namely the propensity towards combining simple verbs with non-verbal elements to create new predicative expressions. This development began with particle+V combinations but spread in Indo-Aryan and Indo-Iranian to encompass adjective and finally noun-plus-verb combinations, due to the combination of factors sketched in the preceding section.

Finally I would like to stress that alongside the kind of static, structural issues such as presenting representations of argument structure, CPs also involve a messier, dynamic dimension of idiomaticity, variation, and partial productivity. Indeed, what at first sight appears to be a fairly uniform phenomenon is far from it, and I have come to mistrust neat and tidy accounts of CPs, which I suspect are in part a consequence of carefully preselected data. The present paper has been an attempt at providing a rough analytical grid, while the leeway within it remains to be plotted in more detailed studies of particular phenomena.

Abbreviations

1/2/3s	first/second/third person singular	M	masculine
A	transitive subject	N	noun
ABS	absolutive	NEG	negative
ART	article	O	object of transitive verb
COMP	complementizer	OBL	oblique
CP	complex predicate	PF	perfect
DEM	demonstrative	PL	plural
ERG	ergative	PPF	pluperfect
F	feminine	PRES	present
FUT	future	PST	past
GEN	genitive	PROG	progressive
GEZ	generic ezafe	PV	preverbal particle
IND	indefinite	Q	question particle
INF	infinitive	REAL	realis
IRR	irrealis	REFL	reflexive
INST	instrumental	S	intransitive subject
L	linker	V	verb
LOC	locative		

References

ACKERMAN, F. & LESOURD, P. (1997): Toward a lexical representation of phrasal predicates, in: ALSINA et al. (eds.), 67–106.

ALSINA, A. (1997): A theory of complex predicates: Evidence from causatives in Bantu and Romance, in: ALSINA et al. (eds), 203–246.

ALSINA, A., BRESNAN, J. & SELLS, P. (eds.) (1997): *Complex predicates*. Stanford: Centre for the study of language and information.

ALSINA, A., BRESNAN, J. & SELLS, P. (1997): Complex predicates: structure and theory, in: ALSINA et al. (eds.), 1–12.

ANNAMALAI, E. & STEEVER, S. (1998): Modern Tamil, in: STEEVER, S. (eds.), *The Dravidian languages*. London: Routledge, 100–128.

BAKER, M. (1988): *Incorporation: A theory of grammatical function changing*. Chicago: Chicago University Press.

BEDIR KHAN, E. D. & LESCOT, R. (1986): *Kurdische Grammatik Kurmancî-Dialekt*. Bonn: Verlag für Kultur und Wissenschaft.

BOSSONG, G. (1985): *Empirische Universalienforschung. Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

BRAUN, F. & HAIG, G. (2000): The noun/adjective distinction in Turkish: An empirical approach, in: KERSLAKE, C. & GÖKSEL, A. (eds.), *Studies on Turkish and Turkic languages. Proceedings of the IXth International Conference on Turkish Linguistics, Oxford, August 1998*. Wiesbaden: Harrassowitz, 85–92.

BURTON-PAGE, J. (1957): Compound and conjunct verbs in Hindi, in: *BSOAS* 19, 469–478.

BUTT, M. (1997): Complex predicates in Urdu, in: ALSINA et al. (eds.), 107–149.

CATTELL, R. (1984): *Composite predicates in English*. New York etc.: Academic.

DEHGHANI, Y. (1998): *A grammar of Iranian Azari including comparisons with Persian*. PhD thesis, La Trobe University, Melbourne.

DIXON, R. (1994): *Ergativity*. Cambridge: Cambridge University Press.

DORLEIJN, M. (1996): *The decay of ergativity in Kurmanci. Language internal or contact induced?* Tilburg: Tilburg University Press.

DORLEIJN, M. & LEEZENBERG, M. (2000): *New registers in modern Kurmanjî: syntactic structures and interference*. Paper held at the First International Conference on Kurdish Linguistics, Kiel, 12–14. May 2000.

DURIE, M. (1997): Grammatical structures in verb serialization, in: ALSINA et al. (eds.), 289–354.

FAIRBANKS, G. & MISRA, B. G. (1966): *Spoken and written Hindi*. Ithaca, New York: Cornell University Press.
GALLMANN, P. (1999): Wortbegriff und Nomen-Verb-Verbindungen, in: *Zeitschrift für Sprachwissenschaft* 18.2, 267–304.
GERDTS, D. (1998): Incorporation, in: SPENCER, A. & ZWICKY, A. (eds.), *The handbook of morphology*. Oxford: Blackwell, 84–100.
GRIMSHAW, J. & MESTER, A. (1988): Light verbs and θ-marking, in: *Linguistic Inquiry* 19.2, 205–232.
HAIG, G. (1998): On the interaction of morphological and syntactic ergativity: Lessons from Kurdish, in: *Lingua* 105, 149–173.
HAIG, G. (2000): *Anatomy of a closed word class: Frequency, regularity and productivity of verbs in Kurdish*. Paper held at the 22. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft, 1–3. März 2000, Phillips-Universität, Marburg.
HAIG, G. (2001). *Zum Verbot der kurdischen Sprache in der Türkei*. Lecture held in the series "Die Kurden: Geschichte – Sprache – Kultur" at the University of Kiel, 13.06.2001.
HAIG, G. (forthcoming): The Corpus of Contemporary Kurdish Newspaper Texts (CCKNT). An exercise in corpus linguistics for modern Kurdish, in: *Kurdische Studien* [for details of the corpus see http://ikarus.pclab-phil.uni-kiel.de/daten/Sprachwiss/CCKNT.htm]
HALE, K. & KEYSER, J. (1997): On the complex nature of simple predicators, in: ALSINA et al. (eds.), 29–65.
HOPPER, P. & THOMPSON, S. (1980): Transitivity in grammar and discourse, in: *Language* 56, 251–299.
JACKENDOFF, R. (1974): A deep structure projection rule, in: *Linguistic Inquiry* 5, 481–505.
KIPARSKY, P. (1997): Remarks on denominal verbs, in: ALSINA et al. (eds.), 473–499.
KRISHNAMURTI, BH. (1998): Telugu, in: STEEVER, S. (eds.), *The Dravidian languages*. London: Routledge, 202–240.
LAZARD, G. (1998): *Actancy*. Berlin: Mouton.
MACKENZIE, D. N. (1961). *Kurdish dialect studies-I*. London: Oxford University Press.
MARGETTS, A. (1999): *Valence and transitivity in Saliba, an Oceanic language of Papua New Guinea*. Nijmegen: Max Planck Institute for Psycholinguistics.
MASICA, C. (1991): *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
MATRAS, Y. (1989): *Probleme der Sprachstandardisierung am Beispiel der Orthographie des Kurdischen*. Unpublished MA-thesis, University of Hamburg.
MATRAS, Y. (1992/1993): Ergativity in Kurmanji (Kurdish). Notes on its distribution and use, in: *Orientalia Suecana* 41–42, 139–154.
MATRAS, Y. (1997): Clause combining, ergativity, and coreferential deletion in Kurmanjî, in: *Studies in Language* 21.3, 613–653.
MEYER-INGWERSEN, J. (1974): Fekr – Ein Beitrag zur Transformationsgrammatik des Neupersischen, in: FRYE, R. (eds.), *Neue Methodologie in der Iranistik*. Wiesbaden: Harrassowitz, 93–305.
MOHANAN, T. (1997): Multidimensionality of representation: NV complex predicates in Hindi, in: ALSINA et al. (eds.), 431–471.
MOSEL, U. & HOVDHAUGEN, E. (1992): *Samoan reference grammar*. Oslo: Scandinavian University Press.
NICHOLS, J. (1975): Ingush transitivization and detransitivization, in: *Proceedings of the Meetings of the Berkeley Linguistics Society* 8, 445–462.
NILSSON, B. (1986): Object incorporation in Turkish, in: AKSU KOÇ, O. & ERGUVANLI TAYLAN, E. (eds.), *Proceedings of the Turkish Linguistics Conference*. Istanbul: Boğaziçi University Publications No. 400, 113–128.
OMAR, F. F. (1992): *Kurdisch-Deutsches Wörterbuch (Kurmancî)*. Berlin: Verlag für Wissenschaft und Bildung.
PAUL, L. (1998): *Zazaki. Grammatik und Versuch einer Dialektologie*. Wiesbaden: Reichert.
RAMM, I. (2000): *Passiv im Kurmancî-Dialekt des Kurdischen*. MA-thesis. University of Kiel.
RITTER, H. (1971): Kurmânci-Texte aus dem Ṭûr'abdîn, in: *Oriens* 21–22, 1–135.
RIZGAR, B. (1993): *Kurdish-English, English-Kurdish dictionary*. London: M. F. Onen.
ROSEN, C. (1997): Auxiliation and serialization: On discerning the difference, in: ALSINA et al. (eds.), 175–202.
ROSEN, S. (1989): Two types of noun incorporation: a lexical analysis, in: *Language* 65, 294–317.
SCHULTZE-BERNDT, E. (2000): *Simple and complex verbs in Jaminjung. A study of event categorisation in an Australian language*. Nijmegen: Max Planck Institute for Psycholinguistics.
SHEINTUCH, G. (1976): Periphrastic verb formation in Persian – a dynamic process, in: CHRISTIE, W. (eds.), *Current progress in historical linguistics. Proceedings of the Second International Conference on Historical Linguistics, Tucson, Arizona, 12–16 January 1976*. Amsterdam: North Holland, 137–146.

SHIBATANI, M. (1973): Semantics of Japanese causativization, in: *Foundations of Language* 9, 327–373.
SKUTNABB-KANGAS, T. & BUCAK, S. (1995): Killing a mother tongue – how the Kurds are deprived of linguistic human rights, in: SKUTNABB-KANGAS, T. & PHILLIPSON, R. (eds.), *Linguistic human rights: Overcoming linguistic discrimination*. Berlin: Mouton, 347–370.
STEEVER, S. (1998): Kannada, in: STEEVER, S. (eds.), *The Dravidian languages*. London: Routledge, 129–157.
TELEGDI, Z. (1955): Beiträge zur historischen Grammatik des Neupersischen. 1. Über die Partikelkomposition im Neupersischen, in: *Acta Linguistica Academiae Scientiarum Hungaricae* 5, 67–183.
VEENSTRA, T. (2000): Verb serialization and object position, in: *Linguistics* 38.5, 867–888.
WATKINS, C. (1964): Preliminaries to the reconstruction of Indo-European sentence structure, in: LUNT, H. (ed.), *Proceedings of the Ninth International Congress of Linguists. Cambridge, Mass., Aug. 27–31*. The Hague: Mouton, 1035–1045.
WINDFUHR, G. (1979): *Persian grammar. History and state of its study*. The Hague: Mouton.

Appendix: Sources of examples

Abbreviation	Source
MEY	Bozarslan, M. Emîn. *Meyro*. Çirok. Borås (Sweden): Invandrarförlaget, undated [collection of Kurdish short stories]
GIR	Cewerî, Firat. 1986. *Girtî*. Sweden. Undated [collection of short stories]
AW*	*Azadîya Welat* [Kurdish weekly newspaper]
ŞK	Şemo, Ereb. 1977. *Şivanê kurd*. Istanbul: Özgürlük [autobiographical novel]
CTV*	CTV [news bulletins published on the internet]
UZUN	Uzun, Mehmed. 1991. *Rojek ji rojên Evdalê Zeynikê*. Stockholm: Weşanên Welat. [novel]
MUŞ	Transcription of spoken Kurdish narratives from the Muş region

* The texts from AW and CTV together comprise the *Corpus of Contemporary Kurdish Newspaper Texts* (CCKNT), described in HAIG (forthcoming).

GEOFFREY HAIG
Seminar für Allgemeine und Vergleichende Sprachwissenschaft
Christian-Albrechts-Universität zu Kiel
Olshausenstraße 40
D-24098 Kiel
GERMANY
haig@email.uni-kiel.de