

Persian Complex Predicates

Julie Ershadi

Swarthmore and Bryn Mawr Colleges

Linguistics Department

jershadi@brynmawr.edu

ABSTRACT. Persian complex predicates are two-part verbal constructions comprised of a non-verbal element and a semantically bleached light verb. A heavy verb such as Persian *xordæn* means ‘to ingest’ (as in food or drink), whereas its light verb counterpart has abstract or bleached semantics; in this case, light verb *xordæn* denotes collision, negative encounters, and other involuntary reception or involvement of an object to a subject. In this paper, we discuss some of the major issues in the field of Persian complex predicates: (1) their compositionality, (2) their idiomaticity, and (3) the hybrid structure or double analyzability of the nominal nonverbal element. We come to these issues with the aim of showing how an approach that focuses on language as a practical tool based on human perspective and experience can yield more accurate interpretations of the various and innumerable linguistic phenomena available for study.

SECTION 1. Introduction.

Persian Complex Predicates (CPrs) are two-part verbal constructions comprised of a non-verbal (NV) element and a light verb (LV). The LV is what is considered a semantically bleached counterpart of a heavy verb (HV) and in a complex predicate construction it accounts for the event semantics and the aspectual properties (Megerdooian 2006). In standard analyses it follows that the NV then selects for the complex predicate’s internal arguments (Toosarvandani 2009, Pantcheva 2008), but there is some controversy over the exact contributions each component of a CPr make to its overall meaning. A heavy verb such as Persian *xordæn* means ‘to ingest’ (as in food or drink), whereas its light verb counterpart has abstract or bleached semantics; in this case, light verb *xordæn* denotes collision, negative encounters, and other involuntary reception or involvement of an object to a subject.

For purposes of manageability, the present study will be restricted to *fodæn*, *kærdæn*, *xordæn*, and *zædæn*, four of the most common light verbs in Persian. Each of these four, glossed with their heavy interpretations, is given in (1).

1. *fodæn* 'to become'
- kærdæn* 'to make/do'
- xordæn* 'to eat'
- zædæn* 'to hit'

The language seems to be moving in the direction of increased reliance on complex predicates, with currently only 115 verbs remaining simple (Mohammad and Karimi 1992). See below for examples, adapted from Megerdooian (2006).

| | | | | |
|---------------------|-------------------|----|------------------------|------------------------|
| <i>agahanidæn</i> | 'inform' | => | <i>agah kærdæn</i> | (informed do/make) |
| <i>piruzinidæn</i> | 'make victorious' | => | <i>piruz gærdandæn</i> | (victorious turn-Caus) |
| <i>aqazidæn</i> | 'begin' | => | <i>aqaz kærdæn</i> | (beginning do/make) |
| <i>peydaginidæn</i> | 'show' | => | <i>neʃan dadæn</i> | (sign give) |
| <i>ayasidæn</i> | 'remember' | => | <i>be yad aværdæn</i> | (to memory bring) |
| <i>zistæn</i> | 'live' | => | <i>zendegi kærdæn</i> | (life do) |
| <i>geristæn</i> | 'cry' | => | <i>gerye kærdæn</i> | (cry do) |

In addition, Persian readily borrows from other languages and forms complex predicates with the new words:

| | | | |
|---------|---------|------------------|-------------------------|
| telefon | kærdæn | (telephone do) | ‘to telephone’ |
| fæks | kærdæn | (fax do) | ‘to fax’ |
| imeyl | zædæn | (email hit) | ‘to email’ |
| klik | kærdæn | (click do) | ‘to click (on a mouse)’ |
| sigar | keʃidæn | (cigarette pull) | ‘to smoke’ |

Persian complex predicates can form between an N, Adj, or Prep (Phrase) and a V
(Karimi 1997):

N + V

| | | | |
|--------|--------|-------------------|------------------|
| dæ'væt | kærdæn | (invitation do) | ‘to invite’ |
| kotæk | zædæn | (beating hit) | ‘to beat’ |
| pænah | bordæn | (refuge carrying) | ‘to take refuge’ |
| atæf | zædæn | (fire hit) | ‘to put on fire’ |

Adj + V

| | | | |
|-------|--------|-------------------|-----------------------|
| bidar | ʃodæn | (awake become) | ‘to wake up’ |
| xærab | kærdæn | (destroyed doing) | ‘to destroy’ |
| sabok | kærdæn | (light do) | ‘to degrade’ |
| pæhn | kærdæn | (wide do) | ‘to spread, to widen’ |

Prep + V

| | | | |
|-----|------------|----------------------|--------------|
| bær | daʃtæn | (upon have) | ‘to pick up’ |
| æz | yad bordæn | (of memory carrying) | ‘to forget’ |

æz sær gereftæn (of head catching) ‘to restart’
 be sær amædæn (to head coming) ‘to expire’

There are formal and informal alternations between complex predicates:

(2) (a) pærvin be æli zæng zæd-Ø.
 Parvin DAT Ali bell hit-3.PAST
 ‘Parvin *telephoned* Ali.’

(2)(a) illustrates usage of a verb meaning ‘to call (on the telephone)’. (2)(b) shows its colloquial replacement, also a complex predicate:

(b) pærvin be æli telefon kærd-Ø.
 Parvin DAT Ali telephone do-3.PAST
 ‘Parvin *called* Ali.’

The present study’s goal is to argue for a revised perspective on Persian complex predicates in order to address two of the major puzzles these constructions present: (1) the various points on the apparent spectrum of compositionality that CPRs seem to occupy and (2) the productive double analyzability of the NV as either complement to or internal to the LV (Megerdooian, 2006; Pantcheva, 2008; Folli et al., 2005). We also introduce for discussion a third puzzle, CPR idiomaticity, not commonly examined in the current literature. The remainder of this study is as follows. Section 2 characterizes the degrees of compositionality CPRs exhibit.

Section 2 also presents data drawn from presenting native Persian speakers with nonexistent, although theoretically possible, original complex predicates. Section 3 relates the discussion to CPr idiomaticity. Section 4 characterizes the dual analyzability of the NV. Section 5 concludes the paper.

SECTION 2. Compositionality and Persian Complex Predicates.

Following a traditional, generative line of thinking, Schnoebelen (2008) discusses non-compositionality in terms of an equation: $z = x + y + n$, where z is the non-compositional meaning, x and y are the components of the form (in the case of Persian complex predicates, the NV and the LV), and n is nebulously termed “something else,” some piece of meaning additional to x and y 's contributions. Meaning_x and meaning_y are associated with the respective components of a form (NV and LV) and meaning_n is noted in a lexical entry. See below:

(3) $x + y$

‘ z ’

Note: n also contributes to z .

In explaining his meaning equation, Schnoebelen clarifies that x and y can have very small values and that in this case n would have a large value to compensate. In linguistic terms, this would mean that x and y contribute very little to z and that most of z 's value comes from an outside or tangential source, n . His example is that of bacon and eggs on a diner's menu: the name of the dish indeed means that upon ordering it, one will receive both bacon and eggs, but it also means that the two will be served fried, or at least cooked (2008:9). While this makes sense,

it still does relatively little to answer the question of where *n* comes from, or how a speaker will know this about the dish called *bacon and eggs*. Common suggestions usually call upon semantic drift, in which case a form or pairing of forms becomes so commonplace that it begins to develop a non-compositional meaning derived pragmatically—bacon and eggs are usually served fried or at least cooked, and the understatement of the fact that they are to be served cooked becomes so accepted as to become an expectation, i.e. part of the phrase *bacon and eggs*—and then the phrase might become lexicalized with some sort of notation keeping track of the manner in which the dish can be expected to be cooked. But a lexicon structured in this way would be overly cumbersome, having to account for and bear the weight of endless information (*n* in Schnoebelen’s notation) connected to specific constructions, even in those cases—which may very well be the majority—where the extra information is easily deduced by other faculties of the mind. Attempting to characterize a grammar and lexicon in which every bit of information about a construction is stored independent of a speaker’s extralinguistic world knowledge is difficult and even somewhat absurd.

Indeed, Lakoff (1977) argued that practical, natural, real-life experiential factors define language’s structure. Language, he says, reflects the way people experience the world. The main claim is this: “A wide variety of experiential factors...determine in large measure, if not totally, universal structural characteristics of language.” This view takes much of the burden off the language faculty and distributes it back among the perceptive faculties, the findings of which it is language’s main purpose to describe. For surely we can say “Colorless green ideas sleep furiously,” and mean more or less nothing by it (*ceteris paribus*), but I classify this speech act with those other myriad, perhaps infinite, things humans are capable of, but which are often considered frivolous, such as whistling, twiddling one’s thumbs, playing, and creating art. From

this perspective, it's easy to see that the famous sentence above is no more meaningless than any of these other acts, all of which are not directly related to eating, sleeping, or reproducing, the things some say are the only real necessities (and thus, speech would follow as a means by which to attain these things, thereby seeming perhaps a fourth, auxiliary necessity), in the sense that no act or thing in this world can truly be devoid of meaning. One can see this in action by writing the famous sentence on the chalkboard in a classroom and asking the students to comment. Sooner rather than later, they begin to search for meaning in the deliberately anomalous sentence. This is the stuff poetry is made of: the human inclination—we might call it a need—to find meaning in things, especially utterances. To argue all this away by saying this analysis takes the famous sentence out of its intended context or domain is to ask oneself to regard language in a vacuum, which is not only an impossible task, but a purposeless one as well. Ask the sociolinguists, for whom a language cannot be understood except on the community level. Devising linguistic analyses based on one's own intuitions about one's own idiolect ultimately gives only an incomplete picture of a language or Language as a whole.

According to Lakoff (1977:239), for the portion of meaning that a listener interprets, "it's not that the meaning of the parts fit together to give the meaning of the whole. Rather, the meanings of the parts mesh with [extralinguistic] knowledge to give rise to the meaning of the whole.... The meaning of the whole is greater than the meaning of the parts." This provides a neat and believable solution to Schnoebelen's equation for non-compositional semantics, $x + y + n = z$, wherein n is a seemingly arbitrary component that adds meaning beyond that provided by x and y , the structural components of a construction. Lakoff put it well when he wrote, "Unless you're a linguist, logician or philosopher of language, it won't come as any shock to you that the

meaning of a sentence can depend on the visual or other experience of the language user, or that the meaning of the whole can be greater than the meanings of the parts” (1977:243).

Following this line of thinking, the analysis of complex predicates seems more manageable. In his 2008 study, Schnoebelen uses the probability that a given English verbal phrase will be broken up, i.e. *I put the book down* vs. *I put down the book*, to measure its degree of compositionality, meaning to what extent the construction’s meaning is a function of the meaning of its parts. This methodology necessarily acknowledges the gradation of compositionality: some forms have transparent meanings wholly comprised of the meanings of their parts and some, on the other side of the spectrum, have an entirely opaque meaning which has no relation to their individual parts. Idioms fall for the most part under this latter end, with some complication arising from overlapping syntax and metaphorical extension. Gries (2002), heavily cited in Schnoebelen’s paper, proposes a paradigm for compositionality. The following is Schnoebelen’s paraphrasing:

Literal: Totally predictable from meaning of the parts: *You can stick the pin in.*

Metaphorical: Not fully predictable from the meaning of its parts because of, say, violations of selectional restrictions that could be accounted for with reference to simple metaphorical or metonymic mappings...or, more importantly, preference violations: *I put down comments.*

Idiomatic: The meaning of the sentence isn’t predictable on the basis of

the parts alone and maximally two simple mappings: *Cerda threw down the gauntlet to Pinochet*.

Thus we have at our disposal the terms *literal*, *metaphorical*, and *idiomatic* to quickly characterize the compositionality of a given form. Forms like *ƿm xordæn* ‘to eat dinner’ are quite *literal*, those like *zæmin xordæn* ‘to fall down’ are *idiomatic*, and those like *særmv xordæn* ‘to catch cold’ are *metaphorical*, or somewhere in between.

But the situation in Persian is more complicated than this. In some sense, the appearances of homophonous LV and HV are distinct in more than levels of opacity. Consider the following example with nominal NVs, adapted from Megerdooian (2006):

(4) NV + HV

| | | | |
|-------|--------|----------------|-------------------|
| qaeza | xordæn | (food eat) | ‘to eat (food)’ |
| xyar | xordæn | (cucumber eat) | ‘to eat cucumber’ |
| sham | xordæn | (dinner eat) | ‘to eat dinner’ |

(5) NV + LV

| | | | |
|----------|--------|-------------------------|------------------|
| kotaek | xordæn | (beating eat/collide) | ‘to be beaten’ |
| faerib | xordæn | (deception eat/collide) | ‘to be deceived’ |
| shekaest | xordæn | (defeat eat/collide) | ‘to be defeated’ |

The forms in (5) appear non-compositional when thinking of their verbal elements as the same as in (4); however, the LV is distinct both semantically and structurally from its HV

counterpart. As Megerdooian (2006) points out and as I have demonstrated with original data below in Section 2.1, native speakers have difficulty naming specifically the meaning of the LV, but that of the HV is readily defined.

SECTION 2.1.

With regards to the following data, we presented a native speaker of Persian with verbal constructions using the verb *xordæn*, some complex predicates and some not, and asked for intuitions about the meaning of (a) the whole construction and then (b) the verbal element alone in each.

(6) *ʃɒm xordæn* (dinner eat)

(a) ‘to eat dinner’

(b) eating

(7) *sib xordæn* (apple eat)

(a) ‘to eat an apple’

(b) eating

(8) *tʃɒp xordæn* (print eat/collide)

(a) Getting printed

(b) ‘the “getting” part/the process of *being* printed’

(9) *særmɒ xordæn*

- (a) ‘getting sick/catching a cold’
 - (b) ‘getting/catching’
- (10) zæmin xordæn
- (a) To fall (on the ground)
 - (b) ‘falling’
- (11) söl xordæn
- (a) “Getting older, years passing you by.”
 - (b) ‘what the person is feeling the years doing’

In each example of a complex predicate, the native speaker defined the LV based on its English counterpart, and in some cases, such as (8) and (11), showed signs of struggle when trying to pin down the meaning contributed by the LV to the predicate. At every turn, the verbal element in a complex predicate can be defined differently, but imagining a lexicon in which this is the case goes against some of the most basic concepts of linguistic inquiry and simply doesn’t make much sense. Instead, many scholars seem to implicitly accept the proposition that light verbs are homophonous counterparts of other, so-called heavy verbs, and thus that the two are distinct (Karimi 1997). The established scholars on the issue seem to assume the LV and HV are a part of the same lexical entry and are distinguished only by their differing argument structure, following Fillmore (1970). According to some views, the LV seems to have more or less no semantic value, and to provide only event and aspectual structure (Folli et. al (2005); Megerdooimian (2006); Pantcheva (2008)). The logical corollary of this is that the NV then holds

all the semantic value of the predicate, but this is unlikely because CPRs allow for stylistic substitution of the LV in formal or written contexts, which has subtle but undeniable effects on meaning and interpretation (see section 4).

Toosarvandani (2009) and Pantcheva (2008) argue that the LV has no semantic value and holds only argument structure and perhaps aspectual information. Many times, *xordæen/zædæen* and *kærdæen/fodæen* form pairs for which the semantic roles are reversed, as in the relationship in English between *buy* and *sell*:

- (12) a. Mærdom (æz dowlæt) færib *xord-æen*.
 People from government deception eat/collide-3.PL
 ‘The people were deceived (by the government).’
- b. Dowlæt mærdom-ro færib *zæd-Ø*.
 Government people-OM deception hit-3.SG
 ‘The government deceived the people.’

Additionally, while the passive voice is arguably nonexistent in Persian (Toosarvandani 2009), the meaning is achieved by verbal alteration, as in below:

- (13) a. Bønu divør-ro pøk *kærd-Ø*.
 Lady wall-OM clean do/make-3.SG
 ‘The lady cleaned the wall.’
- b. Divør pøk *shod-Ø*.
 Wall clean become-3.SG

‘The wall was cleaned. / The wall became clean.’

However, there are other circumstances in which verbal alteration doesn’t yield argument structure reversal:

(14) a. Pesær dʒolu zæd-.

Boy in front hit-3.SG

‘The boy cut in line.’

b. *Mæn æz pesær dʒolu xord-æm.

I from boy in front eat/collide-1sg

(15) a. ʃivɒ bɒlɒ kærd-.

Shiva up do/make-

‘Shiva vomited.’

b. #Qæzɒ bɒlɒ ʃod.

Food up become

Intended: The food was vomited.

Actual: The food was lifted up.

So it seems the alteration is not as productive as first assumed. Again, the situation seems more complicated than that. In order to shed more light on it, we spoke with another native Persian speaker. The experiment devised was as such: present nonexistent, yet theoretically

possible complex predicates and determine (1) how transparent or opaque the guessed-at meanings might be and (2) how productive the argument structure alteration proposed above is.

SECTION 2.2. Original Complex Predicates.

The following data is broken up into four sections, based on whether the NV selected is N, Adj, or Prep. Each NV is paired with *xordæn*, *zædæn*, *kærdæn*, and *fodæn*. If the pairing yields no meaning, it is marked with an asterisk. If the pairing yields an acceptable meaning, that meaning is recorded as a gloss. If by chance a pairing already exists in the language, this is noted in the gloss. Other relevant comments made by the speaker are included in the gloss.

N + V

In (1) below, there appears to be no correlation between the different potential meanings.

1. *kæfsh*, shoe
 - a. *kæfsh xordæn*
shoe eat/collide
'to fall on one's face' (like *zæmin xordæn*)
 - b. *kæfsh zædæn*
shoe hit
'to insult someone'
 - c. **kæfsh kærdæn*
shoe do/make
 - d. *kæfsh fodæn*
shoe become

‘to be struck dumb, to be like a deer in the headlights’

In (2) and (3) below, again there’s no pattern.

2. *dzæbe*, box

a. **dzæbe xordæn*

box eat/collide

b. *dzæbe zædæn*

box hit

‘to sit down and have a conversation’

c. *dzæbe kærdæn*

box do/make

‘to make a box(es)’ (Not an original CPr)

d. **dzæbe fodæn*

box become

3. *fɒdi*, happiness

a. *fɒdi xordæn*

happiness eat/collide

‘to become happy as a result of something external’; *aspectual note*: when using

xordan, the speaker seems skeptical about the subject’s happiness.

b. **fɒdi zadæn*

happiness hit

c. *fɒdi kardæn*

happiness do/make

‘to do a good deed’

d. *ƿædi* shodæn

happiness become

‘to become happy’

4. *omid*, hope

a. **omid* xordæn

hope eat/collide

b. *omid* zadæn

hope hit

‘to give hope’

c. **omid* kardæn

hope do/make

d. **omid* ƿodæn

hope become

5. *sæng*, stone

a. *sæng* xordæn

stone eat/collide

‘to hit the ground’ (like *zamin xordan*)

b. *sæng* zadæn

stone hit

‘to insult (lit., to throw a stone(s))’

c. sæng kardæn

stone do/make

‘to become distant’ (with *xod*, oneself)

d. sæng ƿodæn

stone become

‘to become distant (same as (c) above)’

Adj + V

Of the Adj + V pairings below, only (9)(a) and (b) share alternating argument structure.

The rest are nonsensical, not original, or idiomatic.

6. *bæst*, closed

a. *bæst xordæn

b. *bæst zædæn

c. *bæst kardæn

d. *bæst ƿodæn

7. *kæsif*, dirty

a. *kæsif xordæn

dirty eat/collide

Note: kasif-i xordan is better and would mean to do something depraved

b. *kæsif zadæn

c. kæsif kardæn

dirty make

‘to make dirty’ (not an original CPr)

d. *kæsif* *ƿodæn*

dirty become

‘to become dirty’ (not an original CPr)

8. *bidþr*, awake

a. **bidþr* *xordæn*

awake hit/collide

b. *bidþr* *zædæn*

awake hit

‘to wake up (trans.)’

c. *bidþr* *kardæn*

awake do/make

‘to wake up (trans.)’ (not an original CPr)

d. *bidþr* *ƿodæn*

awake become

‘to wake up (intrans.)’ (not an original CPr)

9. *dʒodþ*, separate

a. *dʒodþ* *xordæn*

separate eat/collide

‘to become separated’

- b. *dzodþ* *zadæn*
 separate hit
 ‘to separate (trans.)’
- c. *dzodþ* *kardæn*
 separate do/make
 ‘to cause to separate’ (not an original CPr)
- d. *dzodþ* *ʃodæn*
 separate become
 ‘to separate (intrans.)’ (not an original CPr)

Prep + V

10. *bþlþ*, above

- a. **bþlþ* *xordæn*
 above eat/collide
- b. *bþlþ* *zædæn*
 above hit
 ‘to hit something high up’
- c. *bþlþ* *kardæn*
 above do/make
 ‘to vomit’ (not an original CPr)
- d. *bþlþ* *ʃodæn*
 above become
 ‘to be lifted up’ (not an original CPr)

11. *zir*, below

- a. *zir* *xordæn*
 below eat/collide
 ‘to be forgotten’
- b. *zir* *zadæn*
 below hit
 ‘to trick someone’
- c. **zir* *kardæn*
 below do/make
- d. **zir* *ƒodæn*
 below become

12. *dzolu*, in front

- a. *dzolu* *xordæn*
 in front eat/collide
 ‘to form a line’
- b. *dzolu* *zædæn*
 in front hit
 ‘to cut in line’ (not an original CPr)
- c. *dzolu* *kærdæn*
 in front do/make
 ‘to boast’

- d. *dzolu ʃodæn
in front become

Again, the first goal of gathering the above data was to examine how transparent or opaque the interpretations might be, and interestingly, a good many of the interpretations above are opaque in their meaning, but there do appear to be more in-between forms, for instance (12)(c) *dzolu kærðæn* (in front do/make) ‘to boast’. However, the argument structure alterations, based on the findings above, seem not to be so productive.

Indeed, the patterns are difficult to identify. The approaches to complex predicates that attempt to locate meaning of the whole as a function of the meaning of the parts come into conflict over how much meaning is in each component: does the NV carry more meaning or does the LV? Does it perhaps shift on a case-by-case basis? These approaches seem erroneous for the simple fact that they require “a complex lexicon with multiple entries, derived by lexical or linking rules” (Megerdoomian 2001). According to Megerdoomian, the complex predicate is instead “formed compositionally by combining the basic components...in syntax,” and “word-formation is not confined to the lexicon.” While this approach neglects to see the lexicon as a system for storing high frequency (usage) constructions as well as idiosyncratic ones, in other regards it makes good sense. Its opposite, the proposal that every single extant complex predicate is stored laundry-list style, has at least three major drawbacks: (1) it does not account for degrees of compositionality found among CPRs, (2) it would require more or less unique definitions for every appearance of a given LV, and (3) it does not account for CPRs’ productivity.

It’s the result of, again, parts constraining but not providing interpretation of the whole. The constraints develop based on usage and frequency: loosely, the more one sees a thing, the

more meaning it carries, and the more specified, and yet abstract, that meaning becomes.

Consider the following scenario.

Kate walks into a classroom on the first day of class. The lid of the trash can in the corner of the room is flipped. If Kate even notices it, she will probably think nothing of it. If it's flipped again the next day, she might notice it again. If it's flipped every day for a month and she notices it every day, when on the thirty-second day it's *not* flipped, she might be surprised, or at least take note of the deviation from (or return to) the norm. Especially if someone else makes note of the situation, as by saying, "The trash can lid is [or isn't] flipped today," Kate might start to wonder. The trash can lid being flipped might come to have meaning, along with the phrase "The trash can lid is flipped." As Gibbs (1993:98) points out, even if Kate and her friend began to use that phrase for meanings beyond or not referring to trash can lids, this would not be a "dead" metaphor (Gibbs' scare quotes), but rather the features of the original event would serve as representatives for some relevant and possibly remote similarity in another event's features. Did the trash can lid mean so many things the first time it was flipped? No, perhaps not. But was it possible that it might mean those things, or have the potential to? Obviously, yes. As Lakoff (1977) says, the "parts constrain, but do not provide, interpretation of the meaning of the whole." This is true not only for non- or semi-compositional phrase-level constructions, but for smaller constructions, like individual words. It also seems like a useful way to conceive of CPRs and of the smaller constructions that make them up.

SECTION 3. Are (Some) Persian Complex Predicates Idioms?

While it's established that complex predicates and idiomaticity are distinct phenomena, the line can often be blurred. Is it that the meaning of the HV is idiomatically warped, or simply that the LV is semantically flexible in a manner that's somehow different from idiomaticity?

There are several reasons why Persian complex predicates might look like idioms. In the first place, their often non-compositional meaning is more or less the classic diagnostic of idiomaticity, but although all idioms are relatively non-literal, not all non-literal constructions are necessarily idioms. Consider, for example, the English words *greenhouse*, *water closet*, and *high five*; and those phrases we consider clichés, as in Jackendoff's (1985) Wheel of Fortune corpus. Consider, similarly, the cross-linguistic phenomenon of metonymy, in which a part of a thing is used to refer to the whole—this process is not totally literal, but every case of metonymy is hardly idiomatic.

How, then, can we decide what's an idiom and what's not? In fact, some scholars seem to consider this issue moot. The need to determine a construction's idiomaticity implicitly regards idioms as a separate, special class, a perspective convincingly refuted in O'Grady (1998) and Jackendoff (1985), among others. A practical component of the argument against is that idiomatic, fixed, and otherwise complex expressions simply make up too much of any given native speaker's repertoire to be reasonably and efficiently relegated to some separate, marginal part of the lexicon, or even outside the lexicon, in what Jackendoff calls "a heterogeneous garbage can". According to O'Grady, the only criterion distinguishing idioms from so-called ordinary constructions is that head licensing specifies down to the terminal node, whereas elsewhere only syntactic categories are licensed. By this argument, idiomatic expressions do indeed behave differently from wholly compositional ones, but they still operate within the same

framework, and, very importantly, they have internal structure. This accounts for many of the strongest controversies in the study of idioms, most notably lexical substitution and varying degrees of compositionality (Jackendoff 1985; Gibbs 1993).

Many studies focus on lexical substitutions—which ones are acceptable and which are not—to determine the nature of an expression’s idiomaticity. O’Grady (1998) distinguishes two types of variation, claiming that *pack a punch/wallop/*slap* is of a different type from *skate on/near/close to/over thin ice*, supporting this assertion with the fact that an American idiom dictionary he cites lists both *pack a punch* and *pack a wallop* but only *skate on thin ice*, and furthermore that *slap* is not acceptable in the *pack a ____* construction but the *skate ____ thin ice* phenomenon “apparently permits substitution of any preposition expressing a spatial relation compatible with the meaning of ice.” However, I’d like to counter this argument—that there exist “standard forms of idioms” that cannot be deviated from—and more generally that idiom breaking is as big a problem as it is characterized to be. For example, I could see the following exchange occurring:

1. A: Boy, I’ll bet the hot sauce in that burrito really packs a punch.
 B: You’d be surprised. I’d say it packs more of a—a slap, maybe, or a love tap.

In the above example, speaker B is clearly making a play on the idiom, possible at least in part because of the particular idiom’s semi-compositional nature. (O’Grady calls it a “post hoc compositional analysis” when speakers identify syntactic overlap once an idiom’s meaning is known.) Speaker B is picking up on the comparison of what we’ll call a *strong effect* and the intensity of a *punch*; by increasing the surface area of the impact in a *slap*, the *strong effect* is

lessened—the hot sauce is not so hot. This undermines O’Grady’s *pack a punch/wallop/*slap* and *skate on/over/near/close to thin ice* distinction and broadens the idea of what constitutes idiomaticity.

Even more easily seen is the comparison below, which makes use of one of the more common idioms in the larger discussion:

2. The shit hit the fan / the crud hit the fan

The second, more polite of these two is perhaps questionable, but easily understood and acceptable. Indeed, putting *the crud hit the fan* into the popular search engine Google yields around half a million results—not too many, but still some. It’s not too farfetched, either, to imagine someone substituting a large number of other words for *shit*, depending on the idea this person wishes to express. It seems that if there exists reasonable cause, all kinds of substitutions can be seen as acceptable. In fact, situations as in (1) and (2) happen many times throughout a speaker’s day. A close listen to the speech of others shows that people break idioms all the time, as much from difficulty remembering the exact wording as from a conscious or subconscious desire to alter the typical meaning. It’s certainly true, of course, that the difference between this fact and the nature of Persian complex predicates is that the substitution permissible for CPRs is much more restricted. As Megerdooian (2006) notes (reproduced in (3) below), the V can be switched out in elevated contexts for synonymous verbs.

3. dɒdæn ‘give’ => bæxʃɪdæn ‘offer’
 ʃodæn ‘become’ => gærdɪdæn ‘turn’

kærdæn ‘do, make’ => nemudæn ‘show’

See Megerdooian (2006) for sentences illustrating the parallel constructions.

Although the above scholars (Jackendoff and O’Grady) resist the conventional attitude toward idioms, they do not deny the *existence* of idioms, and idioms do seem to have some characterizing features that are at least partially unique to other constructions. One feature most agreed on is that idioms are not productive, meaning that speakers cannot or do not create new idioms on the fly, even in a language where new word formation processes (compounding, for instance) are extremely productive. This serves as perhaps the strongest evidence against CPRs as idioms. Returning to Persian, CPRs being productive is perhaps the strongest evidence against the claim for idiomaticity.

A further reason exists to disqualify CPRs from idiomaticity. Complex predicates are structurally distinct from simple predicates in Persian in that for simple predicates, a nonverbal element is an argument of the verb whereas the NV in a complex predicate is a part of the verbal predicate, to be discussed in more detail later in this paper (Megerdooian 2006). This structural distinction is a further and rather strong reason *not* to relegate CPRs *as a whole* to the realm of idioms. The unusual structure of CPRs doesn’t necessarily mean that they are idiomatic. Idioms’ anomalous meaning comes from some extra meaning, perhaps Schnoebelen’s *n*; CPRs’ meaning, however, is a direct result of their structure. It is much more reasonable to regard them as a class of their own, capable alternately of idiomatic interpretation or not, than to conflate their unusual structure with the various inexplicables of idioms. In this view, there is some evidence that totally opaque CPRs might be idioms: for one, they behave similarly in that idiom breaking tends to yield literal interpretation and much the same results from CPR breaking, by which I mean the

unsuccessful substitution of a component of a CPr or other unlicensed modification. In opaque CPrs, the NV, being the provider of the most meaning in the CPr, cannot be substituted for:

4. a. Pærvin *zæmin* xord.
 Parvin earth, ground ate/collided
 ‘Parvin fell.’
- b. #Pærvin *xvk* xord.
 Parvin dirt, earth, dust ate/collided
 Lit. ‘Parvin ate dirt/dust.’
5. a. Pærvin *kpr-ef-o* *tæmum* kærd.
 Parvin work-3.SG.POS-OM whole, all do/make-3.SG
 ‘Parvin finished his work.’
- b. *Pærvin *kpr-ef-o* *hæme* kærd
 Parvin work-3.SG.POS-OM all do/make-3.SG
 Lit. *‘Parvin all did his work.’

In summary, complex predicates seem to have characteristics of their own, whether they be unique to or in addition to idiomaticity.

SECTION 4. Double Analyzability of the Nonverbal Element.

There is some controversy over the relationship between the NV and LV. First, it is important to confront the fact that complex predicates of the form Noun + LV are not always easy to identify in Persian. This is so in part because on the surface they appear to be structurally identical to those simple predicates which are a bare (morphologically unmarked) object and a verb. Their behavior is identical in terms of intonation and stress. Some analyses treat the NV structurally as an internal argument of the LV, but Megerdooian (2006) argues that there is a structural difference between predicates with bare (morphologically uninflected) nominal objects and Complex Predicates, whose NV, she argues, is syntactically linked to the V. She uses syntactical tests such as question formation to support her claim and she uses speakers' intuitions on the semantics of light and heavy verbs to show a difference between complex and simple predicates. Complex predicates and simple predicates with a bare object differ in the type of modification available in each instance. First, although adjectival modification can intervene between both simple and complex predicates, in the first case the adjective modifies the bare object directly and in the second, based on Megerdooian's (2006) analysis, the adjective "behaves as an adverb modifying the whole verbal predicate." The examples below illustrate the difference in interpretation.

1. a. ræft-im ye vyolon-e hes:ɒbi xærid-im
 went-1PL one violin-EZ awesome bought-1PL
 'We went and bought an awesome violin.'

- b. dijæb ye vyolon-e hes:ɒbi zæd-im

last night one violin-EZ awesome hit-1PL

‘We played some awesome violin last night.’

Second, and perhaps more intriguing, the nominal NV is able to take direct object suffixation and be regarded in the discourse as an object of the LV, but not if the NV receives the specific object marker without modification or quantification. The examples in (2) below illustrate this. (2a–b) demonstrate direct object suffixation as it applies to the argument of a simple verb; (3a–b) show acceptable suffixation and (3c–d) show unacceptable structures. A very interesting feature of this phenomenon is that it applies to transparent and opaque, idiomatic CPrs equally, unlike the findings like those in Gibbs (1993) and Schnoebelen (2008) that syntactic flexibility is a function of a construction’s compositionality.

2. a. *dɔneʃju ketɔb xɔnd.*

student book(s) read

‘The student read a book/books.’

b. *dɔneʃju ketɔb-ro xɔnd*

student book-DO read

‘The student read the book (the one previously discussed).’

3. a. *mæn hæʈv æz bæʈʃʃe-hɔ in kotæk-ro xord-æm.*

I even from child-PL this beating-DO ate/collided-1SG

‘I got beaten like that even by children.’

- b. in tæsmim *bozortærin lætme-rov* be dzimnæstik-e irøn
 this decision biggest damage-om to gymnastics-rel Iran
 zæd
 hit
 ‘This decision caused the biggest damage to Iranian gymnastics.’

- c. *mæn hætv æz bætfje-hv *kotæk-ro* xord-æm.
 I even from child-PL beating-DO ate/collided-1SG
 ‘I got the beating even from children.’

- d. *in tæsmim *lætme-rov* be dzimnæstik-e irøn zæd
 this decision damage-DO to gymnastics-EZ Iran hit
 ‘This decision caused the damage to Iranian gymnastics.’

The final contrast to be shown here is question formation, first showing the assertions in (4) that are formed as questions in (5).

4. a. dønɛʃju ketv xvnd.
 student book read-3SG
 ‘The student read a book/books.’
- b. dønɛʃju dærs xvnd.

student lesson read-3SG

‘The student studied.’

5. a. Q: dɒnejju tʃi xɒnd?

student what read

‘What did the student read?’

A: √ Ketɒb.

‘Book/a book.’

b. Q: dɒnejju tʃi xɒnd?

student what read

‘What did the student read?’

A: *Dærs.

lesson

Intended: ‘[Read a] lesson.’ / ‘Studied.’

See Megerdooian’s study for yet more examples like the above. Although she accurately points out and summarizes the major modifications available to CPrs and also extends Folli et al.’s (2005) syntactic analysis of standard CPr structure in a sentence, she does not propose a syntactic description of the available modifications and, furthermore, an exploration of *why* certain modifications are available and others not is neglected. This is not without reason. A

generative framework has no means to describe the factors that motivate the structures it proposes or to even begin to regard the relevance of extralinguistic perceptive faculties to the structure of language (Lakoff 1977). The circular reasoning behind simply creating a rule—revising a hypothesis to reflect a finding and then testing the new hypothesis on the same data that motivated its revision—must come to an end if the serious study of language is to progress.

SECTION 5. Conclusion.

This study's aim is to describe some of the key issues surrounding Persian complex predicates. At its core is the belief that any linguistic phenomenon, including CPs, is a product of human necessity and desire and thus cannot be viewed except in the context of discourse. Persian complex predicates, their idiomaticity, compositionality, and the double analyzability of the nominal NV present challenges to the present accepted linguistic theory and beckon us to widen the perspective of how a language might function.

REFERENCES.

- Aronoff, Mark and Frank Anshen. 1994. "Morphology and the Lexicon: Lexicalization and Productivity". In Andrew Spencer and Arnold M. Zwicky (eds), *The Handbook of Morphology*. Blackwell Publishing, 2001.
- Bird, Brian. 2006. *Tajik Persian Complex Predicates*. University of Oregon.
- Bertinetto, Pier-Marco. 1995. "Compositionality and Non-compositionality in Morphology". In W. Dressler and C. Burani (eds), *Crosslinguistic Approaches to Morphology*. Vienna: Verlag der Osterreichischen Akademie der Wissenschaften, 9-36.

- Fillmore, Charles. 1970. The grammar of *hitting* and *breaking*. In R. Jacobs and P. Rosenbaum, eds., *Readings in English Transformational Grammar*, pp. 120-133. Waltham, MA: Ginn.
- Folli, Raffaella, Heidi Harley, and Simin Karimi. 2005. Determinants of event type in Persian complex predicates. *Lingua* 115:10, 1365-1401.
- Gibbs, Raymond W. 1993. Why Idioms Are Not Dead Metaphors. In *Idioms: Processing, Structure, and Interpretation*, ed. Cristina Cacciari and Patrizia Tabossi, 57-78. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gries, S. T. 2002. Multifactorial analysis in corpus linguistics: A study of particle placement. New York: Continuum International Publishing Group Ltd.
- Karimi, Simin. 1997. Persian Compound Verbs: Compositional or Idiomatic. *Lexicology* 3:2, 273-318.
- Irnex Technologies. *Farsi Dictionary*. Copyright 2010. [9/10/2010]. Available from <http://www.farsidic.com>.
- Jackendoff, Ray. 1985. The Boundaries of the Lexicon. In *Idioms: Structural and Psychological Perspectives*, ed. Martin Everaert, Erik-Jan van der Linden, Andre Schenk, and Rob Schreuder, 133-166. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lakoff, George. 1977. Linguistic gestalts. *CLS* 13:236-87
- Megerdooonian, Karine. 2006. The Status of the Nominal in Persian Complex Predicates. *NLLT*.
- Mohammad, Jan and Simin Karimi. 1992. "Light" verbs are taking over: Complex verbs in Persian. In *Proceedings of The Western Conference on Linguistics (WECOL)*, vol. 5, pp. 195-212. Department of Linguistics at California State University, Fresno.

- Pantcheva, Marina. 2008. Noun preverbs in Persian Complex Predicates. *Tromso Working Papers on Language & Linguistics: Nordlyd 35*, special issue on Complex Predication, ed. Peter Svenonious and Inna Tolskaya, pp. 19-45. CASTL, Tromso.
- Schnoebelen, Tyler. 2008. Measuring the compositionality of phrasal verbs. *3rd Workshop on "Quantitative Investigations in Theoretical Linguistics"*. Helsinki.
- Toosarvandani, Maziar. 2009. Ellipsis in Farsi Complex Predicates. *Syntax* 12:1, 60-92.
Blackwell Publishing Ltd.