

Definiteness Effect in Complex Predicates of Modern Farsi

Mohammad-Mehdi Vahedi

University of Ottawa

This paper examines the argument structure and formation of complex verbal structures known as compound verbs, or Complex Predicates (CP) in Modern Farsi (MF) which show properties of X^0 *morphological objects* and *phonological words* as well as X^{max} *phrasal predicates*. To solve the dichotomy in their behaviour and to account for their formation, I argue that the CPs count as non-listed phrasal lexical integers formed within V-bar in MF, and V-bar has a special, idiomatic status. On the other hand, I argue for a syntactic level of argument structure and an organized and rich level of lexical component that exists alongside and parallel to the syntactic component so that it interacts with the syntactic component at the different stages of syntactic derivation and allows word formation to take place after and during the syntactic derivation. The wordhood of CPs is accounted for by assuming that the syntactic, X-bar configuration of CPs serve as input to the conflation and incorporation at the lexical component or morphological structure.

1-Introduction

In this paper I address the argument structure and syntactic formation of complex verbal structures known as compound verbs, or Complex Predicates (CP) (Cattell 1984), in the linguistic literature of Modern Farsi (MF), or Persian, the official language of Iran. CPs in MF are verbal complexes consisting of a first, non-verbal element or preverb (PV) and a Light Verb (LV). The PV may be a noun, adjective, verbal participle, adverb, or a whole PP. The Two form a single semantic, syntactic, morphological and phonological unit corresponding to simple verbs in Persian or other languages.

Contrary to the English CPs that are normally discontinuous, the CPs in Persian are continuous units in that the combination of the PVs and the LVs form a phonological word with a single stress pattern attracted to the PV. That is the PV destresses the verb which normally falls on the main verb (Ferguson 1957).

However, CPs seem to introduce a major dichotomy in the language in that when they are used in their primary function as the finite verbal predicates of a tensed clause, their pieces are freely separable and the CPs lose their *phonological, and morphological wordhood*. On the other hand, CPs may freely undergo frequent derivational and inflectional processes that apply to simple verbal stems, and to lexically derived words/items in the word-formation domain or the morphological structure (MS).

As far as the first property is concerned, I argue that CPs are formed within V-bar and make up non-listed phrasal integers in surface sentential syntax (s-syntax), or the computational component (cf. Chomsky 1993) (section 7). In section 7-8, I account for the word-hood property of CPs by arguing that the s-syntactic, X-bar structure of CPs may be

carried down to their MS, or may be isomorphic to their MS. As a result, the wordhood of CPs in MF falls under the rubric of noun or complement incorporation or in the lexical domain or MS where a syntactic complement forms a single phonological and morphological word with the LV while still retaining its syntactic structure. Section (2) provides a descriptive analysis of the morphological properties of CPs. Section (3) examines the syntactic properties of CPs as finite verbal predicates. Section (4) gives a theoretical background. Section (5) accounts for the semantic relationship between the LVs and PVs as well as their syntactic configuration within V-bar. Section (6) argues for an abstract weak objective case for the nominal PVs that is concomitant with a weak existential reading on them.

2-The Data & Wordhood of CPs: Complex Predicates as Morphological Objects

CPs in Persian are verbal constellations consisting of a first, non-verbal element or preverb (PV) plus a bleached, Light Verb (LV), or a so-called full, thematic verb when its thematic meaning is reduced, pushed back and backgrounded. Some examples, in Masdar or Infinitive form¹, of the first type with LVs are given in (1-6), and of the second type in (7-8). Many of the CPs with LVs come in contrastive pairs of causative/accomplishment vs unaccusative/inchoative. This is reflected in the (b) examples in (1-6):

- | | | | |
|-----|--------------------|-------------------------|--------------------------|
| 1.a | 'āšofte=kard-an | agitated.ppr=do-INF | to make agitated |
| .b | 'āšofte=šod-an | agitated.ppr=become-INF | to become agitated |
| 2.a | sard=kard-an | cold=do-INF | to make cold |
| .b | sard=šod-an | cold=become-INF | to become cold |
| 3.a | 'az dast=dād-an | from hand give-INF | To give up, lose |
| .b | 'az dast=raft-an | from hand go-INF | To get lost |
| 4.a | šafā=dād-an | cure=give-INF | to cure |
| .b | šafā=yāft-an | cure=find-INF | to be cured |
| 5.a | 'az kār='endāxt-an | from work=throw-INF | to make disabled/useless |
| .b | 'az kār='oftād-an | from work=fall-INF | to be disabled |

¹By masdar and/or infinitive I mean the derived forms of a simple (or compound) verb which are formed by affixing **-an** to the (past) verbal stem as in (1-8). So by CP-INFinitive I mean the masdar or the infinitive from of a CP. Masdars are taken to be deverbal nominals and have a dual function as the gerunds of English.

6.a	čub=zad-an	stick=hit-INF	to punish
.b	čub=xord-an	stick=collide-INF	to be ounished

The examples (7-8) show that CPs are also formed with full, thematic verbs, in addition to LVs in the above examples. Both sets of CPs may undergo different morphological derivations and processes applicable to single, simple verbal heads and words in Persian.

7.a	qazā=xord-an	food=eat.ps-INF	to food-eat, to eat
.b	'āb=xord-an	water=eat.ps-INF	to drink (water)
8.a	ketāb=xānd-an	book=read.ps-INF	to book=read, to study
.b	dars=xānd-an	lesson=read.ps-INF	to study

In (1) a verbal (past) participle, in (2) an adjective, in (3) a whole prepositional phrase, P+noun (NP), in (4) a predicative nominal, in (5) a P+NP, in (6) a common instrumental noun, and in (7-8) some common nouns are used as PVs².

All CPs (potentially) undergo a level of nominalization resulting in their infinitive/masdar form (CP-INF) ending in *-an* (1-8). CP-INFs and simple infinitives (Masdars) have a listed sense and are used as the lexical entries for verbal predicates in MF. However, contrary to the English CPs which are normally discontinuous units (Cattell 1984), the infinitive form of Persian CPs (CP--INFs) functions as single *phonological words* in the sense that the whole CP-INF has a single word stress which is attracted to the PV, and behave as *morphological objects* in the sense of DS&W (1987) in that it behaves like a single, simple noun/word and shows syntactic atomicity.

Nothing can intervene between the PV and LV of a CP-INF. CP-INFs in (9) function like simple nominal heads and may appear as a subject (9a), DO (9b), object of a preposition (9c) or other positions sanctioned for simple lexical nouns, i.e., syntactic head positions where heads, X⁰, project to XP levels.

²Sign = is used to relate a PV to a LV, and the status of the CP as a morphological unit, while - sign before an element shows other derivational relations between items in this paper. The other abbreviations used include PV for preverb; ppr for past participle; ps for past stem; EZ for Ezāfe morpheme; S for singular; RA for the strong DO and strong ojective case; MS for morphological structure and word-formation domain; NEG for negation; DO for direct object or direct internal argument; IND for indicative; PL for plural; INF for infinitive; LV for light verb; IDF for indefinite; PRG for progressive.

- 9.a **bus=kard-an / (busid-an)** nešāne-y-e mohabbat 'ast
kiss-do-INF / (kiss-INF) sign-EZ affection is
Kissing is a sign of affection.
- .b Mehri **qazā=pox̄t-an** ro dust=na-dār-e
Mehri food=cook-INF RA love=NEG-have-3S
Mehri does not like cooking.
- .c mo'allemin bāyad 'az **čub=zad-an** bačče-hā xod-dāri kon-and
teachers must from stick=beat-INF children self-have do-3PL
The teachers must abstain from beating the children.

However, all different types of CPs in MF may undergo numerous other derivational and inflectional processes normally applicable to simple verbal stems and to synthetic compounds. In (10), we see cases of deverbal adjectives by affixing *-ān* to the present stem of a CP:

- 10.a **čort=zan-ān** "while taking a nap" .b **sorfe=kon-ān** "while coughing"
.c **qadam=zan-ān** "while walking"

The structure in (10a) can only be [[**čort=zan**] -ān] with *-ān* added to the present stem of the CP **čort=zan-**, since the verbal part, i.e., ***zan-ān**, is ungrammatical by itself. The same is true for (10b-c).

CP-INFs and other morphological derivations of CPs count as syntactic atoms and the whole domain of the CP/compound exhibits what has been known as *Anaphoric Island Constraint* (Postal 1969) in that the (nominal) PV is non-referential and may not have a referential relationship to the elements outside the compound. Thus, in (11), the equivalent of (9b), and (12) the nominal PVs of the CP-INFs are inaccessible to anaphoric relations:

11. *Mehri **qazā_i=pox̄t-an** ro dust=na-dār-e vali Amir **'un_i=pox̄t-an** ro dus=dār-e
Mehri food=cook-INF RA love=NEG-have-3S but Amir it=cook RA love=have-3S
Mehri does not like cooking, but Amir likes it-cook.

12. *man 'az **sigār_i=kešid-an** bad-am mi-yā-d, va hargez **'un_i-(o)=ne-mi-keš-am**
I from cigarette=smoke-INF bad-my come-3S and never that-(RA) NEG-IND-smoke-1S
I hate smoking cigarettes and will never smoke it.

One may conclude with Sproat (1988) that an anaphoric relation is a property of maximal projections and the PVs in the above CPs being X^0 level items are not subject to this constraint. If Shibatani & Kageyama (1988:474) are correct in claiming that Anaphoric Island Constraint is "not a constraint on words in general, but only on lexically derived

words" then (11-12) show that CP-INFs (and other forms derived from the CPs through morphological affixation) have the property of lexically derived words. To conclude CPs in MF behave as *morphological objects*. The combination is that of [_V⁰ X⁰+V⁰] level formations which then undergoes morphological processes like simple verbal stems.

3-Syntactic Properties of CPs:Complex Predicates as Syntactic Phrases

The main function of the CPs is in the syntactic computational component where they act as the main finite predicate of tensed clauses. A cursory survey of the behaviour of CPs as finite predicates in syntax shows that they violate the lexical integrity and syntactic atomicity observed above, and function as phrasal units, i.e., maximal projections. That is, finite CPs show syntactic properties opposite to what is expected if CPs are *morphological objects*. PVs can be separated from LVs by finite modal-like verbs (13a-b), PP complements and modifiers (13d), relative clauses (13e), and modifiers (13c). PVs can, in general, be gapped and coordinated (13f-g). That is, the finite use of CPs in syntax shows every sign of maximal projections or XPs.

- 13.a. Morad zamin= xāhad xord
 Morad earth will.3S hit.ps
 Morad will fall down.
- .b. ?'unhā raqs= dār-an(d) mi-kon-an(d)
 they dance have(PRG)-3PL IND-do-3PL
 They are dancing.
- .c. yek zamin-e saxt-i xord-am
 one ground-EZ severe-IDF collide-1S
 I fell down severely.
- .d. 'unhā 'estāqbāl-e garm-i 'az Amir kard-and
 they welcome-EZ warm-IDF from Amir did-3PL
 They gave Amir a warm welcome.
- .e. 'agar zamin-i ke man xord-am to mi-xord-i mi-mord-i
 If ground-IDF that I hit-1S you IND-hit-2S IND-die-2S
 If you fell down the way I did, you would die.
- .f. mina 'az man 'ejāze=gereft vali mahtab --- =na-gereft
 Mina from me permission got.3S but Mahtab --- NEG-got.3S
 Mina got permission from me but Mahtab did not.

.g 'Estefigraf' xub šoru' (=...) va bad tamām=kard
 Stefigraf well start... and bad end=did.3S
 The Stefigraf (a soccer player) started well and ended up badly.

The results of this section show that PVs in CP-structures are maximal projections. These observations contradict the syntactic atomicity and the lexical integrity we noticed with respect to CP-INFs in the previous section. Finite CPs are phrasal predicates and I argue that they form non-listed *generalized lexical integers* (Szabolsci 1984, 1986) within V-bar and have a structure like [_V XP+V⁰].

4- Theoretical Background

We see that CPs in MF exhibit properties of both X⁰ morphological objects, and syntactic maximal phrases, XPs. That is, CPs in Persian show properties similar to what Ackerman (1995:3) claims for the phrasal predicates in Hungarian:

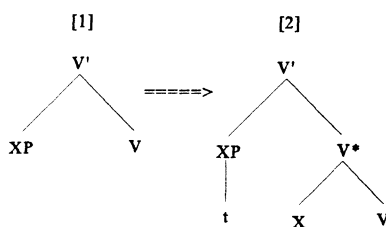
PV and V exhibit the sort of lexical information and co-occurrence conditions ordinarily associated with (derived or compound) words and participate in derivational morphology, but their status as morphological objects, i.e. phonologically integrated and syntactically atomic, is questionable, since their pieces are separable in syntax.

Diagrams [1] and [2] correspond to the possible structures, and representations of the CPs in MF in section (3) and (2) respectively. That is, if we assume that all types of PVs count as subcategorized, (first) sister complements of the head of LVs at the syntactic component, then the two form a syntactic and semantic unit within V-bar which corresponds to the finite phrasal CPs in section 3. On the other hand, if we define incorporation as the phenomenon of (the head of) a sister complement/object combining with a verb to form a single morphological unit, while still retaining its syntactic status, then the diagram in [2] which is derived through head movement of the PV-complements to the head of the LV (Baker 1988) is a potentially possible representation of the morphological behaviour of the CPs in Persian as observed in section 2 (cf. Sadock 1991; Mohanan 1995 for a different interpretation of incorporation that excludes the head movement shown in [2], as well as Ghomeshi and Massam 1994).

However, the crucial question to answer is the relationship between the two representations. Is [1] derived from [2], or is [2] derived from [1]. Or else, are they separate, independent structures and representations somehow related to each other.

Suppose that we assume that

complex head in [2] is initially formed in the lexical and word-formation component, or MS, according to the principles applicable at that level (cf. Di Sciullo & Willaims 1987) and the PV and LV form a tight morphological and phonological unit, and a complex head (§



2). However, when the complex head in

[2] is inserted/projected in/to X-bar syntactic derivation as a finite verbal predicate, then the morphological and phonological tie, i.e., lexical integrity, between the two is freely broken and the V⁰ head in [2] suddenly converts to a phrasal unit, V-bar, like [1] (§ 3).

However, this would mean that when an X⁰ lexicalized, or derived, synthetic head/CP, like [2], is projected or inserted in syntax, it loses its lexical integrity and functions as an X^{max} phrasal complex. This is a phenomenon that is highly improbable and totally unattested, to my knowledge. Thus, I take it for granted that the derivation of [2] cannot have preceded [1], and [2] cannot have served as an input to derive [1]. Furthermore, as we will note presently, almost all conditions, and constraints for the formation of CPs in Persian are syntactic, semantic, and pragmatic/discourse phenomena pertinent to the surface sentential syntax (s-syntax henceforth)³.

Rather, I argue that it is the other way round. CPs are syntactic products and are formed in the syntactic computational component, i.e., s-syntax. Their particular X-bar configuration within V-bar, [1], counts as a non-listed *generalized lexical integer* (Szabolsci 1984, 1986), or a phrasal CP, that is subject to the *predication condition* (Higginbotham 1987).

Almost every V-bar in the syntactic component, [1], corresponds to a (potential)

³ I use s-syntax as equivalent to X-bar syntax and syntactic computational component in the more recent changes in the GB theory (Chomsky 1993, 1994)

conflated verbal head, V^0 , in the lexical component, or MS, [2]. The two domains might be co-present (though more work need to be done), and show isomorphic syntactic configurations and relationships between the verbal heads and the PVs. The two levels are formed according to similar syntactic, X-bar principles (§ 6). That is, the syntactic representation in [1] is available, and is isomorphic, at the lexical component, and/or MS as well which then leads to lexicalization, and wordhood of CPs at that level [2], and § 2. That is, word-formation [2] may and can follow the syntactic derivation of CPs [1]. Yet, in other words, [1] may serve as input to derive [2].

5- Semantic Analysis

In this section I examine the semantic relationship between the LVs and the PVs in persian, and the syntactic properties of V-var as a non-listed phrasal CP.

Nominal PVs in CP-structures count as weak NPs in the theory of Milsark (1974, 1977) where the strength of a noun is determined by the strength of its determiner system which in turn characterizes the noun as weak or strong. Weak nouns have an existential interpretation and count as cardinality words as opposed to strong nouns (characterized by strong determiners) which are quantificational. Only weak nouns may appear in existential constructions. Weak determiners are argued to be intersective or symmetric in that they denote a relation between two sets and only take the cardinality of the intersection of those two sets into account, and disregard the other members of the domain. As a result they can only appear in existential constructions that introduce a new set, already unavailable in the domain of discourse. That is, existential constructions assert or introduce the existence or availability of a new set, and hence may only take weak nouns that have an existential interpretation in their post-verbal position (Milsark) or NP-in-VP position (Safir (1985; 1987)) as in (14a). (14a) is true if the set denoted by the *cats* is not an empty set. Strong NPs on the other hand cannot appear in existential constructions since their interpretation would lead to a contradiction or tautology given the existential interpretation induced by standard existential sentences (14b). In (14b) the strong partitive reading induced by *some of* is incompatible with the assertion of existence provided with the *there ... is* construction. *Some of the cats* in (14b) not only takes the cardinality of the intersection of two sets into consideration but also says something about the set of cats that are not in the garden, i.e., that

it cannot be empty. That is, from (14b) it is induced that only some of the cats in the domain of discourse are in the garden and the rest are not in the garden. However, (14a) disregards this latter notion, as given in (15a-b) corresponding to (14a-b).

- 14.a There are some cats in the garden.
 .b *There are some of the cats in the garden.

- 15.a $\|some\|_{exist}(N) = (X \subseteq E: |N \cap X| \geq 2)$
 .b $\|some\|_{part}(N) = (X \subseteq E: |N \cap X| \geq 2 \ \& \ |N - X| \neq \emptyset)$

Milsark explains the Definiteness Restriction (DR) by considering the formative "there" uniformly as a marker of existential quantification, analogue to the operator $\exists$ in predicate logic. It quantifies over the cardinality of the set denoted by the weak indefinite NPs, and cannot take already (universally) quantified NPs/determiners as arguments because this would lead to double quantification on the NP and is hence ruled out as ill-formed or anomalous.

Safir (1985) dispenses with the existential nature of "there be" constructions but keeps Milsark's weak vs. strong distinction. Safir (1985) and Reuland (1983) account for DR in existential sentences by proposing an unbalanced theta chain between *there* in subject position and the indefinite NP in the postverbal position (or VP-in-VP). For Safir the chain is motivated by the case theory that requires the indefinite NP to receive case. The indefinite NP is generated in a non-case theta position, while the formative "there", or an equivalent expletive *pro*, in subject position receives NOM case and by coindexation transfers it to the indefinite NP which lacks case. For Reuland (1983) the unbalanced theta-chain is motivated by theta-theory that requires the *there* in subject position to be coindexed with the theta-marked indefinite NP in postverbal position in order to pass the Extended Projection Principle of Chomsky (1981) at LF. For both of them the NP coindexed with *there* in the subject position must be weak and indefinite since co-indexation of *there* with an R-expression (=strong NP) would violate the condition C of binding theory. Weak, indefinite NPs, however, do not count as R-expressions and hence are exempt from the binding condition C. They, thus, account for DR in existential constructions. The identical property of Milsark, and Safir and Reuland is that they all require an expletive element in the subject position of the existential sentences, and the verb of such constructions must be either "be" or an intransitive or "inside verb", i.e., equivalent to an unaccusative verb.

The relevance of the DR for CP-structures of MF is that LVs, similar to "there be" and the inside verbs of Milsark and Safir, seem to induce a DR on their VP internal position in that only a weak (predicative), existentially quantified NP may occur in the VP-internal position of these verbs as in (16a) and (17a). (MF is SOV and the VP-internal position in MF is equivalent to post-verbal NP position or VP-in -NP position in English that these authors talk about.)

16.a Amir **šenā** **kard**
 Amir swimming did.ps.3S
 Amir did swimming/swam.

.b *Amir **šenā-ro** **kard**.⁴
 Amir swimming-RA did.ps.3S

17.a Mina **šafā** **peydā=kard** / **yāft**
 Mina cure visible-do.ps-3S / find.ps-3S
 Mina was cured. Lit. Mina found cure.

.b *mina **šafā-š-o** **peydā=kard** / **yāft**
 Mina cure-her-RA visible-do.ps-3S / find.ps-3S
 Mina found her cure.

Determiners are used infrequently in MF and the difference between weak and strong NPs in VP-internal (object) position is determined by the type of case, syntactic position, scrambling and the postposition **rā/ro/-o**. Strong referential and quantified NPs in the object position (specific NPs) are followed by the post-position **rā/ro** while weak object NPs (non-specific, indefinite NPs) are not overtly (case) marked. In (16-17) the main LVs, i.e., **kard-an** "to do", and **peydā=kard-an** "to find/yāft" "to find", are transitive but cannot take a strong, presuppositional NP as their DOs (16b-17b) and may only take a weak, non-specific object (16a-17a). The weak DOs in (16a-17a) clearly have an existential interpretation. In (16a) we talk of the occurrence of an event denoted by the nominal **šenā** "swimming", the name of the action. The verb is an abstract, empty verb which is unspecified inherently and hence cannot take a specified, strong complement in its object position (16b). Similarly, in (17) the main predicate is compatible only with a weak object with an existential

⁴ I take the postposition **rā/ro/-o** as a sign of a strong, quantificational object NP with a presupposed referent, as well as the sign of strong objective case on the NP. This is opposed to its lack on an object that indicates a weak, existential, predicative NP with weak objective case (Vahedi 1996)

interpretation (17a), and not with a strong NP in spite of the fact that predication holds. In (17a) the subject Mina comes to be in a state that was non-existent or unavailable to her before.

However, there is a main difference between (16-17) and standard existential sentences (14-15). The subject position of (16-17) is filled with a lexical NP while in the standard existential sentences the subject is an overt or covert expletive which seems to be an indispensable part of these theories. My analysis departs from the standard ones in that I claim, similar to Szabolsci (1984, 1986), that the Definiteness Effect (DE) is not a property of *there* sentences with an inside, intransitive verb or *be* as the main verb, rather DE is the property of NP-in-VP position or VP-internal object position of a class of verbs that I will call, following Szabolsci, as *bleached existential predicates*.

Existential predicates are empty or bleached verbs, and their semantic, non-logical, or thematic content/meaning is either totally bleached and empty, or is reduced and backgrounded in certain uses. Their meaning is reduced to the assertion of EXISTENCE, availability or prevalence of an action or trait characterized by the indefinite NP, or a (predicative) XP. The existential predicates may contain any of the primary logical components EXIST, BECOME, BRING into EXISTENCE, DO, MAKE, BE, GO, COME to BE in the sense of Dowty (1979). Many of such predicates also contain a Motion, Change of Location/State component as part of their logical meaning. On the other hand, their non-logical semantic substance/meaning is either totally bleached or is backgrounded. They, however, keep an impoverished argument structure. These predicates induce an existential reading within their VP-internal position, just like standard *there be* existential constructions, and as a result are only compatible with a weak, existential NP, or a predicative XP (PP, Adj, Adv).

Notice the following Hungarian examples (18) from Szabolsci (1986) who identifies DE in the object position of certain transitive and intransitive sentences with lexical theta-marked NOM-subjects, and hence no expletive element, overt or covert, in subject position. Szabolsci (1984) calls them "bleached existential predicates". This observation undermines the claims of the standard analyses of existential constructions with respect to the existential nature of "there" (Milsark, Safir, Reuland). In (18a-b) a transitive verb with a lexical subject shows DR on the object position in Hungarian. She claims that in (18a) the existence of *two*

pens/some milk is the result of what Mary did. The verb only conveys an existential reading, hence a definite NP in object position leads to ungrammaticality (18b). On the other hand, (18c) with the perfective prefix *meg-* means that Mary successfully located *two pens*, the existence of those pens is independent of whether Mary found them. The object may be specific or non-specific, and no DR is observed.

- 18.a Mari talált tollat / két tollat / (némi) tejet.
Mary found pen-ACC two-ACC some milk-ACC
Mary found a pen / two pens / some/sm milk.
- .b *Mari talált(a) a tollat / P. tollat / minden tollat.
Mary found the pen-ACC P's pen-ACC every pen-ACC
Mary found the pen / Peter's pen / every pen.
- .c Mari megtalált (a) két tollat / a tollat / ...
Mary PERF-found two pen-ACC the pen-ACC
Mary found two pens / the pen / ...

Szabolsci (1984; 1986) argues that in (18a), and in a list of DE (light) verbs that she provides the "specific content of the verb that triggers the definiteness effect is backgrounded; meaning is reduced to the assertion of (a change in the state of) existence" which she refers to as "bleached" verbs. The addition of the perfective prefix *meg-* to these verbs brings about a fully specific meaning and no longer asserts (a change in the state of) existence, and hence no DR. She calls them "full" verbs.

These observations are true for English as well, where the same surface verb can be used in the bleached sense and in the full sense, which makes it more difficult to spot DE, and requires more attention to the meaning of the sentence, while in Hungarian the prefix *meg-* reduces the distinction between the two types of sentences to plain ungrammaticality (Szabolsci 1986):

- 19.a John gave Mary a/*the hard time. (Szabolsci (1986, her 9))
- .b John gave a/the bottle/*the hard time to Mary.

(19a) which shows DR would be expressed without the perfective prefix *meg-* in Hungarian, while (19b), which does not show DR, would be expressed with the prefix *meg-*. Notice that in (19a) with the DR, the specific non-logical content of "give" is backgrounded, i.e., "give" does not express the notion of "transference of an object" (Jackendoff 1990), rather it expresses a context in which Mary is caused to change to a new situation which she was not

NPs (and other compatible XPs) can appear in this position which then form a CP within V-bar with the LV as explicated below.

Partee (1987) argues that each syntactic category corresponds to a family of semantic types as opposed to just one single type in the Montagovian approach. She distinguishes three basic types of noun phrase (NP) interpretations: Referring ($\langle e \rangle$), Predicative ($\langle e, t \rangle$), and Quantificational NPs ($\langle \langle e, t \rangle, t \rangle$). She assumes, following Partee & Rooth (1983), that each category is lexically assigned the simplest unmarked type to capture its meaning, and proposes a set of general type-shifting principles that allows shifting from one type to another for the members of the same category, thus accounting for the different interpretations of NPs in different contexts. It is more natural to treat proper names and singular pronouns with the basic referential type $\langle e \rangle$, but a proper name like *John* has to be lifted to the type $\lambda P[P(j)]$, generalized quantifier type, a set of sets of individuals $\langle \langle e, t \rangle, t \rangle$ in order to interpret the conjunction *John and every woman*, while common and instrumental nouns are of the simple semantic type predicative $\langle e, t \rangle$.

I suggest that the nominal PVs in the CP-structures of MF (also Hungarian) are lexically of the semantic type predicative, and the existential reading is a property of predicative nominals of the type $\langle e, t \rangle$ Higginbotham (1987). The distribution of predicative nominals is constrained to the VP-internal position of those verbs that can induce a weak existential reading, or DR, in their VP-internal position, that is, bleached predicates of existence or LVs and backgrounded verbs. This also takes care of the fact that PVs in MF may also be adjectives, PPs, and adverbs that are uncontroversially of the predicative type in syntax.

On the other hand Higginbotham (1987), Safir (1987), and Williams (1994), on quite independent theoretical grounds, argue that indefinite NPs in the standard existential constructions are not arguments, rather they are *predicates* and their distribution and function is characterized by the *predication condition* (Rothstein (1983)) that requires that every predicate must be saturated in syntax. Higginbotham (1987:46) in accord with Rothstein claims:

- 21.a All arguments are saturated.
- .b All predicates are unsaturated.

Saturated and unsaturated are used to mark a distinction between phrases that denote

objects/entities, and those that have one or more open positions into which objects are to go in order to saturate them. Higginbotham (1987; 1985) claims that predicate NPs in the existential constructions have an Event position in their argument structure that enters into a *theta-identification relation*, a kind of modification relation, with the Event position of the main verb of the clause. Hence that the E position of the weak predicate noun is co-indexed and identified with the E position of the existential verb, and the former thus modifies the latter and counts as part of the projection of the latter in syntax.

I suggest that *theta-identification* is the type of thematic and semantic relation between all types of PVs and LVs in the CP-structures of MF in finite clauses. Nominal, adjectival, adverbial and PP PVs in CP-structures function as VP-internal complements of the LVs or bleached existential predicates, and satisfy the subcategorization requirement of the LVs for an internal argument. Yet PVs are not true arguments since they are not referential $\langle e \rangle$, nor quantificational $\langle e, t \rangle, \langle t \rangle$; rather they are predicative $\langle e, t \rangle$ and function as predicate modifiers $\langle e, t \rangle, \langle e, t \rangle$. Predicate modification requires a close knit between the modifier and the modifiee which I suggest to take place within the deepest level of V-bar, or the iterated VP-shell in the sense of Larson (1988). The two elements now function as a single phrasal predicate and V-bar counts as a phrasal CP in finite clauses. Following Szabolsci (1984; 1986) I suggest that the predicate modification within V-bar produces what she calls *lexical integers*.

Before explaining this notion let us see where the CPs get their meaning or substance from. We claimed that LVs lack non-logical substance and are bleached, but we know that CPs as a whole are not bleached and do have lexical, non-logical substance. LVs are substantiated by their predicative complements or the PVs:

22 **Substantiation:**

- (i) Every predicate of natural language must have some non-logical content. Therefore,
- (ii) If the meaning of a predicate contains at most logical constants and variables, it must enter into a "closest possible" syntactic relation with something whose meaning (also) contains some non-logical constant. (Szabolsci 1986:11)

V-bar is the locus of substantiation, a kind of predicate modification in syntax and produces a *lexical integer*. Szabolsci claims that all DE-complements of the LVs incorporate into the V-bar in Hungarian, and function as verbal modifiers. V-bar functions as a "lexical integer" in Hungarian, and, I claim, in MF as well.

23 Lexical Integers:

Lexical integers are syntactic products which can be "theta-role assigners, but no theta-role assignment takes place, or needs to take place, within them", Szabolsci (1984, 1986), [underlining mine].

This implies that LVs do not assign theta roles, even the quasi theta-role of Chomsky (1981). Furthermore, the move divorces subcategorization from theta-marking, and reduces the import of theta-criterion as an independent well-formed condition on syntactic (Deep-)structures. Predicate modifiers are partially incorporated into the LVs within the V-bar, and count as the subcategorized complements of the LVs, and/or their substantiators, but not as true arguments. No theta-role assignment takes place within V-bar, rather V-bar functions as a theta role assigner for elements outside V-bar in the sentence, i.e. a lexical integer (23), (for more detail see Szabolsci 1984, 1986; Vahedi 1996; and also Mohammad & Karimi 1992).

In sum, in this section I have shown that CPs are syntactic, phrasal lexical integers formed within V-bar, as in [1], in MF. Before providing an analysis for the lexical, wordhood properties of CPs I will briefly explain the case-marking of nominal PVs within phrasal CPs.

6- Case of Nominal PVs

Let us study the case and structural position of the nominal PV in (24) (the causative counterpart of (17a) above):

24. Pezešk Mina rā šāfā dād
Physician Mina RA cure give.ps.3S
The physician cured Mina.

The strong, referential DO, Mina, is followed by rā. However, the nominal PV šāfā "cure" in (24) must also receive case to pass case filter. We seem to have an instance of double object construction in (24). Notice the following example of double accusative construction in German, and its English gloss, from De Hoop (1989):

25. ...dass er mich Deutsch gelehrt hat
that he me-ACC German-ACC taught has
"...that he taught me German."

Mich receives structural accusative case and can function as the subject of the passive

sentence for (25), while Deutsch, in spite of being marked with ACC case-marking desinence, must have an inherent case since it cannot function as the subject of a passive sentence for (25) (see de Hoop 1989; Larson 1988; Belletti 1988). The same has been proposed for the English gloss of (25) by Chomsky (1981), and Larson (1988) (also Belletti 1988) who state that me has structural and German inherent case. Thus, in the English gloss for (25), the double object construction, only me can function as the subject of the passive sentence *I was taught German*, and not *German*, **German was taught me*. According to Larson (1988), the NP *German*, in English gloss (25), is an "adjunct argument" of the verb and the verb forms a CP *teach German* which then satisfies its other open positions in syntax.

This is in accord with our analysis for (24) where *šāfā* "cure" is a predicate (non-argument). However, I disagree that *Deutsch* "German" and *šāfā* receive inherent case which is fixed with a specific theta-role. *Dād-an* "give" in (24) and *teach* in (25) are bleached and backgrounded verbs and cannot assign theta-roles since they lack any non-logical, thematic or semantic substance (see Szabolsci 1984, 1986; Vahedi 1996 for argument that LVs lack the potential for theta-role assignment, and argument against theta-role transfer between LVs and PVs as in Grimshaw & Mester 1986). Furthermore, the complex verbal predicates *šāfā=dād-an* "to cure" and *Deutsch=gelehrt* "teach German" count as *Lexical Integers* (CPs), and hence no theta roles are assigned, or need to be assigned, within them. The nouns *šāfā* and *Deutsch* "german" do not function as arguments, rather are (adjunct) predicates, so by the definition of the lexical integers (23) they do not need a theta role even if the main verbs were theta-role assigning verbs. Also note that *šāfā* "cure" in (24) and *Deutsch* "German" in (25) and the English gloss occupy identical syntactic positions with respect to the verb, i.e., the deepest level of VP-projection within V-bar and cannot be reordered.

I agree with Jaeggli (1986), Siewierska (1984), De Hoop (1992), and Mohammad & Karimi (1992) that Deutsch "German" in (25) and *šāfā* in (24) are assigned structural objective cases directly by the verb within V-bar in VP-internal position. This case is assigned to non-argument, weak NPs and is different from structural (ACCusative) case which is directly checked/assigned by the verb when the latter raises (overt or covert) to the head of functional projections (AgrO) above VP as the case of *Mich* "me" in (25).

To sum, I have shown that CPs count as phrasal predicates called *lexical integers*

in the syntactic computational component, throughout to the spell-out and PF representation. There is a very close syntactic and thematic relation between the PVs and the LVs since they are thematically and semantically as well as syntactically knitted together within V-bar in MF. V-bar is the deepest level of verb projection in the sense of Larson (1988) and H&K (1991, 1993) where the verb/LV takes the PV as its *single internal complement* to the exclusion of the other elements within VP as in the case of double object constructions like *teach German* in (25), and in the dative construction *he taught German to me* where *taught to me* counts as a CP to the exclusion of *German* (see Larson 1988; H&K 1991, 1993; and Vahedi 1996).

7- Conflation or Incorporation in Lexicon as a method of lexicalization

To this point I have tried to account for the syntactic, phrasal properties of the CPs as XPs. I have to show now how the lexicalized CPs which behave as "Morphological Objects" are formed. To do so, I adopt the level of lexical representation, or argument structure (and/or morphological structure) called Lexical Relational Structure (LRS) in the theory of Hale & Keyser (H&K) (1988, 1991, 1993, 1994).

7.1- The Lexical Component

I take the lexical component as a highly organized and rich level of semantic, lexical, and phonological information and structure. In particular, I assume that the lexical component is composed of Lexical Conceptual Structure (LCS), a syntactic level of argument structure called Lexical Relational Structure (LRS) (H&K 1991, 1993), as in [1-4], and lexical phonological properties/formations, as well as the idio-syncratic properties of lexical items. I take the morphological structure (MS) or word-formation domain to be a sub-component of the LRS with syntactic, X-bar properties. Morphological word-formation builds on the basic syntactic structures of the LRS and the derivations formed from the LRS representations, [1-2].

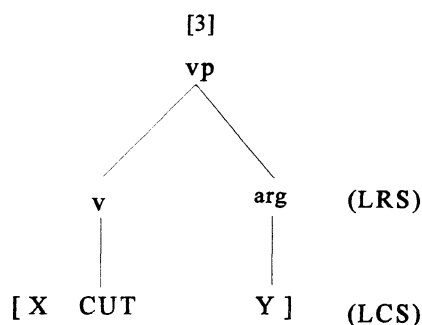
The lexical and/or morphological component is present parallel to, and stretches along the syntactic derivation so that multiple interactions between the two are possible (Parallel morphology of Borer 1988, 1991; and Jackendoff 1994). This theory allows word-formation to take place during and after the different stages of the syntactic derivations and computations, i.e., morphology after syntax. While I do not specifically argue for *late*

insertion (Hale & Marantz 1993, 1994) the model adopted in this paper allows syntactic structure to be carried down to a post-syntactic morphological structure.

I also assume that the grammar is made up of a syntactic, computational component that draws from the lexical component at different stages of syntactic derivation to form structures determined according to the core syntactic, X-bar theory and UG. Multiple interactions between the two levels are possible at different stages of derivation.

7.2- Lexical Relational Structure (LRS); Universal Categorical Features

LRS is a level of argument structure or Lexical Structure. It is "an abstract syntactic projection of the lexical categories, embodying the basic syntactic organization of its arguments" defined by X-bar theory in the lexical component. The LCS and lexical structure of the transitive verb *cut* is given in [1] from H&K (1986:23):



[1] shows the way the arguments

of the predicate *cut* are represented in the lexical structure, i.e., LRS/argument structure. It is isomorphic to s-syntactic X-bar projection of *cut* in that the verbal head takes an argument/complement which it c-commands. An important difference with Predicate Argument Structure (PAS) is that the subject/ external argument of *cut* [the variable X] is not represented in lexical structure [1] at all, since it is external to the verb phrase. Subject is a property of finite clauses at s-syntax required by predication theory and is subcategorized by INFL.

LRS representations take the notational form of syntactic trees defined by X-bar theory. According to H&K (1991:13) "Like d-structure, LRS representations involve the notions category, projection, government (and complement), adjunction, and specifier." The major lexical categories corresponding to the traditional parts of speech V, A, N, P, as well as certain functional categories like the Kase phrase (Fukui 1986), are also involved in the

LRS representation. In other words LRS representations import "most aspects of syntax in the more general sense", and are thus subject to general principles of syntax. That is, the major lexical categories share their basic, x-bar categorial features and structures at s-syntax and at the lexical component.

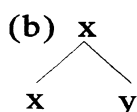
The lexical categories define a system of relations between a head and its arguments and specifiers (if any). Heads project their categorial features. Lexical categories define a *universal class*, and are

divided into some notional, conceptual types. Verbs are of the notional type (dynamic) 'event'. Events are not predicates at LRS. The subject is a requirement of predication. Verbs not being predicates at LRS, they do not have a subject at this level, meaning that they do not project a Spec as in (4b). Nominals are of the notional type 'instances', or 'entities'. They simply project a head, extending to a maximal projection, without a Spec, or

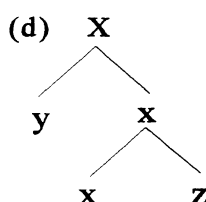
complement (4a). Prepositions are interrelational. They define a relation between a complement and a subject, that is they project a complement and a Spec at LRS (4d). Adjectives are attributive, attributing a property to something, hence are inherently predicates. They project a subject, Spec, but no complement at LRS (4c).⁵

[4]

(a) x



(c) y / x



⁵The categorial features/properties of lexical items are assumed to be part of our intuitive, innate capacity of language and UG and hence is identical to the different levels of grammar like LRS, PF, LF and computational component or s-syntax (H&K 1991, 1993; Chomsky 1993). That is, the basic X-bar, categorial features of the lexical items are isomorphic at all these levels. However, structures, and derivations may be built on them according to rules and principles specified to each level.

The s-syntactic and LRS representation of the *universal lexical categories*, which are assumed to be primitives of the theory, is given in [4] (H&K (1994: 25, (58)). In [4], *x* stands for the head of a primitive lexical category, and *y*, and *z* correspond to its arguments. Lexical items of the primitive universal category *noun* are shown by [4a]; of the category *verb* by [4b]; of the category *adjective* by [4c]; and of the category *preposition* by [4d].

7.3- The Single Complement Hypothesis

The X-bar categorial feature for the *universal category verbs* in [4b] does not make any distinctions between *full lexical verbs* vs *light verbs*, nor does it differentiate between monadic unergative verbs,

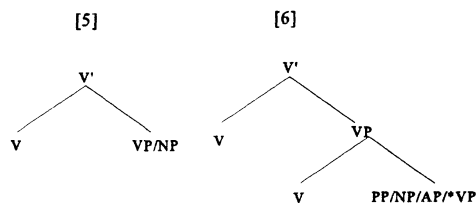
that are assumed to have a sole external argument with no internal argument like *laugh*, unaccusative verbs, with a sole internal argument and no external argument like *reach* and *sink*, simple activity transitive verbs with an external and an internal argument like *do*, *have*, and triadic verbs like *give*. All members of the *universal category verb* dominate a sole sister complement and no external argument, i.e., verbs have no specifiers or subjects at the level of LRS representation since they are not predicates at that level. [4b] is repeated in [5].

[5], with the NP-complement, corresponds to the LRS for unergative verbs (26a) (cf.[8-11]), simple transitives like the LV *kard-an* "to do" in MF, its stylistic equivalents, and to verbs that function as a simple transitive verb (26b).

26.a The child laughed.

.b We had a good laugh.

If the complement in [5] is a VP then it can expand to a configuration like [6], a causative complementation structure. The higher *V'* is a causative verb and the inner VP is an inchoative/unaccusative verb of change of state, location, etc. The complement of the inner VP can only be a N, A, or PP but never another VP since this would lead to double causativization, which is generally undesirable. This is the type of relational structure which



is employed in the causative use of the LVs **kard-an** "to do", and *suru* "to do" in Japanese (Grimshaw & Mester 1986), **dād-an** "to give" in MF and English which regularly alternate with some corresponding inchoative verbs shown by the inner VP in [6], ergative alternation verbs like *to thin*, *to break*, dative constructions with *to give* as in Larson (1988), and *to put*, and location and locatum verbs like *to shelve*, *to jail*, *to saddle*, de-adjectival verbs, like *to redden*, *to flatten* in English⁶.

I suggest [5-6] as the basic syntactic representation, i.e., the basic l-syntactic (LRS) and s-syntactic configuration, for all types of CPs, causative or inchoative, in MF. [5-6] count as the syntactic representation for both thematic, full (causative and inchoative) verbs as well as their corresponding LVs in CP-structures. The difference between full vs light verbs arises in the s-syntax representation. It is the choice of a full vs a LV, and the semantic properties of the single internal complement of the verb which leads to conflation and formation of CPs at LRS and substantiation at s-syntax.

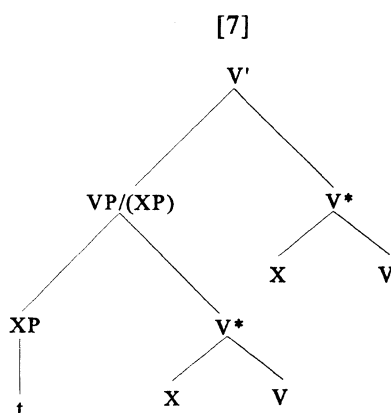
LVs (and CPs) share the basic structures [5-6] at both levels of LRS and syntactic computation throughout to the spell-out. We see that the two levels are isomorphic and the PVs in s-syntax occupy the same syntactic position with respect to the LVs as they do at the l-syntax/LRS. That is, the PVs count as the single subcategorized complements of the LVs and are c-commanded by the head of the LVs.

However, the head of XP, or the head of the PP, or the small clause may conflate

⁶ During the syntactic derivation and computation (s-syntax) verbs behave as predicates and require a subject/specifier because they occur in a temporal context and must receive a temporal reference (H&K 1993, 1994). The temporal reference is structurally reflected by the functional projections (within IP) over the VP. The verb occurs within the TP to receive a T(ense)-value. Given that verbs inherently lack a subject/specifier position [4b], I believe that subjects are not base-generated within VP at the syntactic computation. I assume with Chomsky (1993) that specifiers belong to the residue domain of a verb at syntax, but follow recent work by Ghomeshi (1996) who suggests that subject of a verb occurs at the specifier of a P(redication) P(hrase) above VP in Persian. Lack of functional features and checking at LRS would make the appearance of the specifier positions redundant at LRS.

into the head of the lower and/or the higher Verb to form contrastive accomplishment and inchoative CPs at LRS, [7], given the semantic conditions and syntactic configurations are met throughout the syntactic derivation (§ 5-7). The only difference is that no such conflation takes place at s-syntax and the XP remains in its original position within V-bar. Any complements of the XP (=PV) will count as the complements of the whole CP and will remain within the XP at LRS, and may be invisible. At s-syntax these complements behave as the structural subjects, DOs, PPs of

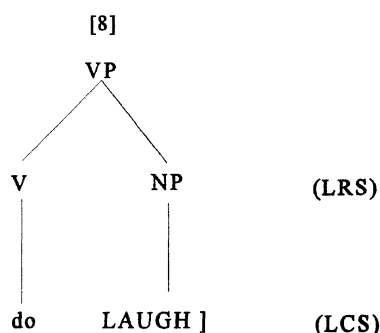
the sentence and occupy the GF positions provided by the PVs, LVs, and INFL (the topic of the following sections). The conflated structure at LRS for CPs corresponding to [6] is shown in [7] without any further discussion.⁷



7.4 Unergatives

The major repercussion of this theory is that all the members of the lexical category verb, including the monadic, unergative verbs, i.e.,

laugh, *sleep*, *jump*, etc., must also



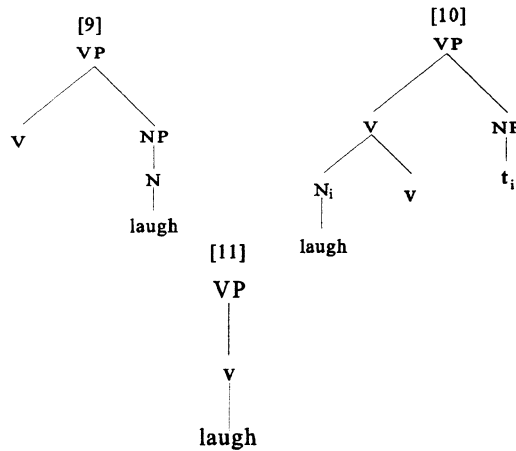
⁷ The XP shown in the parentheses, adjacent to VP in [7], indicates that the higher verb, V', may simply select an NP/XP as a complement, and not an embedded VP as in [5]. In this case no (lower) inchoative event, and no causative/inchoative contrastive CPs, would be available. The verb counts as a simple transitive one and the resulting CP would count as an unbound, atelic, activity verb, and/or an unergative CP like *do+laugh* in MF and Basque as in [8-11].

take a complement but no subject/ external argument at LRS. H&K (1988; 1991) notice that in Basque the equivalent of the simple, unergative verbs like English *laugh* is expressed by a transitive LV, like English *do*, and the nominal *laugh*, e.g., *laugh+do*, just like *xande=kard-an* "laugh (N) + do; to laugh" in MF. They take this observation to propose a LRS identical to the LRS of the transitive *cut* [3] for all unergative verbs (cross-linguistically) as in [8]. The LCS and LRS of the Basque *laugh+do*, roughly corresponding to English simple unergative *laugh*, (and to *xande=kard-an* "laugh+do, to laugh" in Persian (27b) below) from H&K (1991: 56, (31)), and H&K (1988: 48, (11)) is given in [8].

In LCS [8] the standard external argument variable, *x*, is not represented at LRS, because an external argument is not "committed" to an internal grammatical function position in the lexicon/LRS, and only projects to a syntactic subject position in combination with INFL. LAUGH, the "restricted" variable, corresponds to the relevant action and functions as a tacit, implicit "object", or the name of the action, of the empty verb *do*. It is equivalent to the "unrestricted" variable *Y* in [3].

The LRS of the simple, unergative *laugh* in English is given in [9-10-11] from H&K (1991:31). They argue that the actual LRS of a simple unergative verb, like *laugh* in English, is derived through a pattern of lexicalization, called "conflation" by Talmy (1985), a variant of head-to-head movement, or incorporation at Lexical-structure/LRS (or lexical-syntax by now), which

derives [10] from the initial verb-object structure [9]. [10] is licit in that it observes all conditions of head-to-head movement, and the ECP (Travis 1984; Baker 1988).



Notice that [9] is analogous to [8] of Persian and Basque with the difference that in the latter the head of the V is filled with a LV like *do*, but in [9] it is an abstract empty verb (V). In [10] the head of the DO, in [9], conflates, or incorporates into the abstract empty verb and forms a complex verbal head, a single lexical unit corresponding to English *laugh*.

The full derivation of unergative verbs, [10], is visible/available in lexical-syntax, or LRS, and forms part of our "integral and accessible" knowledge of lexical entries. However, it is not clear, rather is doubtful, that the full derivation with the traces are accessible/visible at s(entential)-syntax as well. They thus suggest [11] as the s-syntactic representation of *laugh* before combining with INFL (H&K 1991:37).

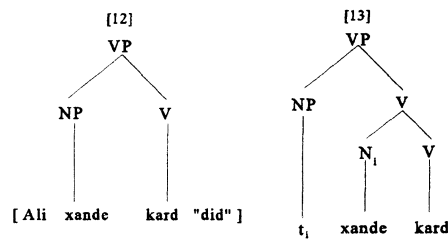
The morphological wordhood of the CPs in Persian (and in Basque) can be accounted for by assuming that the (nominal) PV, conflates into the verbal head in an instance of head-to-head movement, or incorporation, and forms a complex lexical head as in [8] at LRS. No such conflation occurs in s-syntax, hence the corresponding phrasal CPs.

In Persian, the equivalent of both the simple unergative verb *laugh* of English, i.e., **xandid-an** "to laugh" (27a), and the CP *do=laugh* corresponding to Basque, i.e., **xande=kard-an** "laugh=do, to laugh" (27b), are available. The PV in (27b) exhibits phrasal properties since it can be freely separated from the LV, and the whole CP violates the lexical integrity and syntactic atomicity attested of lexically-formed words. The PV may be modified, relativized, etc., as we noticed in section (3).

- 27.a Ali **xand-id**
 Ali laugh.3S "Ali laughed."
- .b Ali **xande = kard**
 Ali laughter did.3S "Ali laughed"

Given this theory of LRS representation and the analysis proposed above for the derivation of English *laugh* in [9-11] and [8] above, I suggest that the derivation of the simple verb **xand-id-an** "to laugh" in (27a) is exactly as given in [9-11] where the nominal **xande** "laugh, laughter" as the head of the complement of an abstract empty V conflates into the V and then fuses with the past tense suffix *-id*. Since successive vowels are prohibited within the same syllable in MF, and no vowel actually begins a syllable, the final vowel of nominal complement **xande**, following the familiar trend in the phonology of MF, is deleted and we get the past stem **xand-id** in (27a).

The finite phrasal CP in (27b) is formed in the syntactic domain within V-bar with the properties of lexical integers as in [8 & 12]. However, recall that this CP may also undergo different derivational (and inflectional) processes and



exhibits wordhood properties as discussed in section (2), i.e., **xand-e=kon-ān** "laughing", **xande=kard-an** "to laugh". Perhaps the interesting point is that an identical derivation just like [9-10] and (27a) can also account for the wordhood of the CP in (27b), with the difference that this time the head of the empty verb is phonologically realized as the LV **kard-** "do" in MF just like [8 & 12]. Similar to [10], the head of the "restricted" nominal variable/complement, i.e., **xande**, conflates into the head of the LV **kard-**, forming the complex verbal head/CP **xande=kard-** "do+ laugh" at the level of LRS as shown in [12-13]. This complex head now has properties of morphologically-derived words, but formed according to the principles of syntax, hence may be also called a *syntactic word* (Di Sciullo & Williams 1987).

We note that the lexical structure of the CP **xande=kard-an** in [13] is identical in every respect to that of the simple verb **xandid-an** and **laugh** in MF and English in [9-11] which accounts for their identical semantic and syntactic properties.

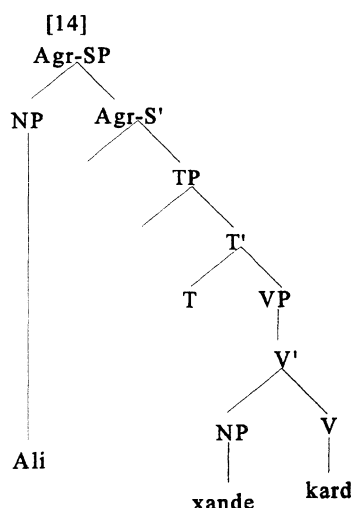
It seems more interesting to note that the syntactic representation and derivation of the finite, phrasal CP in (27b) is quite isomorphic to its lexical representation [8 & 12] as shown in [14]. In [14] the restricted variable **xande** functions as the *single complement*, and the implicit object of the LV **kard-** within V-bar. The two form a tight semantic, thematic, and syntactic unit within V-bar and enter into theta-identification relation which is a joint projection of the thematic and syntactic properties of the two (Higginbotham 1985). However, since no incorporation or conflation has taken place in s-syntax (computational component), the two retain their independent syntactic identity resulting in the type of phrasal properties we noted in section (3). The (nominal) PV (=the restricted variable) remains within the deepest level of verb projection, i.e., the Nuclear Scope, and is closed by

the default existential closure operator, hence the existential reading that is always concomitant with the semantic interpretation of such PVs (Diesing 1992; Heim 1982).

In this theory the morphological formations may take place at the different stages of, and after, the syntactic derivation and interacts with the latter repeatedly. Basic phrasal syntactic structures like [4, 8 & 12] are isomorphic at the s-syntax and lexical component (LRS and MS) and may serve as input to head-movement and morphological derivations. It is properly capable of accounting for the dual phrasal and wordhood properties of CPs of MF. I assume that the Morphological Structure (MS) counts as one of the subcomponents of LRS with syntactic, X-bar properties. The categorial features/properties of lexical items, [4], are assumed to be part of our intuitive, innate knowledge of language and UG and hence is identical to the different levels of grammar like LRS, PF, LF and computational component or s-syntax. As a result no extra burden is laid upon the minds of the native speakers to account for their structures and formations.

8. More Evidence from CP-Structures

In this section the formation of an alternating pair of recursive (double) CPs with the causative *dād-an* "to give" and its inchoative alternant *gereft-an* "to get" that reflect the structure [6-7] is analyzed where these verbs act as LVs and take a nominal PV, i.e., *qarār* "placing, setting" as their single internal complements leading to the CPs *qarār=dād* "to place, to set", and *qarār=gereft-an* "to be placed, be soothed". However, the interesting point is that the nominal PV, *qarār* "placing, setting", has its own independent set of arguments, LRS, and selects an interrelational PP complement, which then makes up a case of a CP with recursive PVs as in (28-29) below:



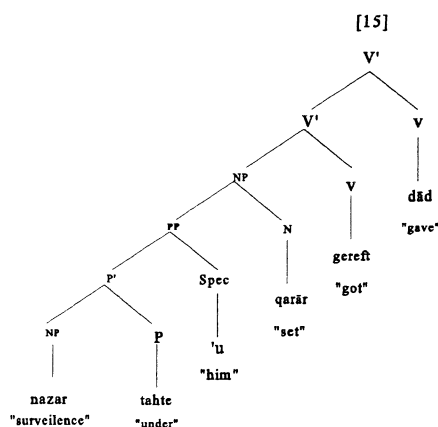
28. Polis 'u-rā [_v taht-e nazar qarār dād]
 police him-RA under-EZ surveillance setting give.ps.3S
 The police put him under surveillance.

29. 'u [_v taht-e nazar qarār gereft-an]
 he under-EZ surveillance setting get.ps.3S
 He was put under the surveillance.

Not only the nominal PV **qarār** but also its P+NP complement satisfies the syntactic and semantic requirements for PVs in sections (2-3) and hence form CPs with the above two alternating verbs. The LRS [15] shows the relationships among the different elements of the sentences in (28-29) before the process of conflation.

The higher V' corresponds to the causative LV **dād-an** "to give" and the lower V' to the unaccusative LV **gereft-an** "to get". Neither verb has a Spec (subject or external argument) position. The noun **qarār** "setting" functions as the single complement of the inner VP in [15]. The nominal in turn selects an (interrelational) PP as its single complement, [PP 'u taht-e nazar] "he under surveillance" headed by the preposition **taht** (-e) "under". Both the nominal PV **qarār** "setting", and the NP complement of the P, i.e., **nazar** "surveillance", count semantically as predicative and non-arguments. On syntactic grounds they count as part of the V-bar and hence form a syntactic *lexical integer* or phrasal CP by substantiating the LVs.

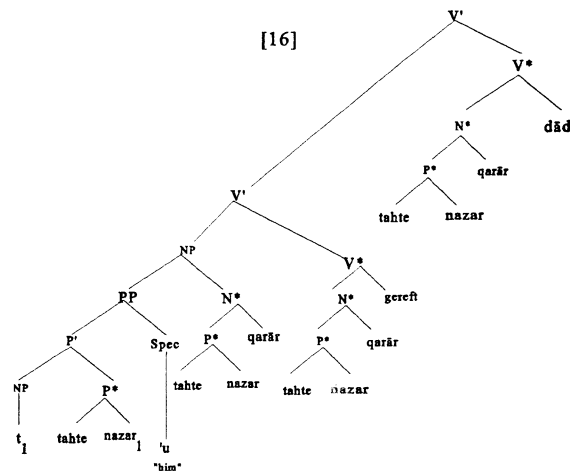
On the other hand, their LRS [15] has every syntactic property required for conflation. The LRS after conflation resulting in a single lexical head is shown in [16]. In [15-16] **taht-e** "under" is the head of the PP selected by **qarār** "setting". The complement of the preposition, i.e. **nazar** "surveillance" conflates into its head P forming P*. This complex in turn conflates into its own head, **qarār**, skipping the Spec position,



forming N*. This complex then incorporates into any of the two dominating LVs to form the CPs in (28-29). Notice that in no case can the Spec position be conflated.

The subject/Spec of the PP, 'u "him", functions as the subject of the unaccusative LV in (29), the only desirable situation. The subject of an unaccusative verb can only occupy a VP-internal position in a syntactic configuration, as 'u "him/he" does in [15-16]. 'U also functions as the DO of the causative LV in (28), a perfect causative/ inchoative or ergative alternation. The conflated heads in [16] now behave as morphological objects, formed according to the principles of syntax, hence *syntactic words*. They may enter into various derivational processes including CP-INFs, i.e., *taht-e nazar=qarār=dād* "to put under surveillance", and *taht-e nazar=qarār=gereft-an* "to be under surveillance".

The configuration in [15] and the finite phrasal verbs in (28-29) form a generalized lexical integer when they occur within the functional categories through the derivation in the computational component. The pronoun 'u "him" raises out of the V'



in [15] to function either as the subject of the DO of the causative CP in (28) or the subject of the unaccusative CP in (29). What remains within V' is shown in (28-29), within the square brackets, and functions as phrasal verbal predicate in MF.

9. Concluding Remarks

CP-formation is a syntactic process of compounding in MF. Although the combination of preverbal elements (PV) and Light Verbs (LV) exhibit the sort of lexical

information and co-occurrence conditions ordinarily associated with (derived or compound) words and participate in derivational morphology, their status as syntactically atomic morphological objects is questionable: the pieces of CPs are separable in syntax when they function as finite verbs/ predicates.

I have dealt with this double nature by showing that CP-formation in Persian is a productive syntactic process. Semantically, LVs induce existential interpretation within V-bar; the PVs are predicative with a weak existential reading and enter into Predicate Modification with the LVs (Higginbotham 1987). Syntactically, the predicative PVs substantiate the LVs within V-bar resulting in phrasal CPs which have some properties of lexically formed words called non-listed lexical integers. These are subject to the predication condition. At the lexical level, on the other hand, the categorial features of lexical items are represented in the format of X-bar theory in a way that is isomorphic to their syntactic categorial representation. This level is called Lexical Relational Structure (LRS): it is a lexical-syntactic level of argument structure as in H&K (1991, 1993). The X-bar configuration of the PVs with respect to the LVs at LRS can lead to conflation and grammaticalization of the CPs in the lexicon and the formation of complex synthetic words that have properties of syntactic and morphological objects. I have shown that both full verbs and their equivalent LVs have identical argument structures. They both select a single complement in the format of X-bar theory. I have also defended a morphological structure that allows word-formation to take place alongside, and after the different stages of syntactic derivations.

In this system of CP-formation, and grammar in general, the "computational system draws from the lexicon to form derivations, presenting items from the lexicon" in a format determined by UG (Chomsky 1993). This format is already shared by the lexical component and is accessible for the computational system, i.e., a core syntax in the format of X-bar theory that reflects the lexical and categorial features and properties of lexical items [4]. There is multiple interfaces and interactions between the syntactic derivation and lexicon or LRS representations which includes the Morphological Structure (MS) where word formation, compounding and conflation take place. The LRS representations and Morphological Structure seem to stretch along and go parallel to the syntactic computational derivation so that word-formation may take place at any stage of the syntactic derivation to

spell-out as in the theory of Distributed Morphology of Halle & Marantz (1993), and Parallel Morphology of Borer (1988). CP-formation is a process from syntax to the lexicon not a projection from the lexicon to the syntax. Only if the formal semantic and syntactic requirements and configurations are satisfied during the syntactic derivation, will there be a lexical level of conflation and grammaticalization as well.

Selected Bibliography

- Ackerman, F. (1995) Systemic Patterns and Lexical Representations: Analytic Morphological Words. To appear in J. Kenesei (ed.) *The structure of Hungarian*, Vol. 5, JATE/SZEGEO.
- Baker, M. (1988) Incorporation: A Theory of Grammatical Function Changing. University of Chicago Press, Chicago.
- Baker, M. (1985) Syntactic Affixation and English Gerunds. WCCFL 4, 1-11.
- Belletti, A. (1988) The Case of Unaccusatives. LI 19, 1-35.
- Chomsky, N. (1993) A Minimalist Program for Linguistic Theory. In Hale, K. & J. Keyser (eds.) *The View from Building 20*. The MIT Press, Cambridge.
- Chomsky, N. (1994) Bare Phrase Structure. ms. MIT.
- Diesing, M. (1992) Indefinites. Cambridge, MA: The MIT Press.
- Di Sciullo, A.M. & Williams, E. (1987) On the Definition of Word. *Linguistic Inquiry* Monograph 14, MIT Press, Cambridge.
- Dowty, D. (1979) Word Meaning and Montague Grammar, Reidel, Dordrecht.
- Enç, M. (1991) The Semantics of Specificity. LI 22, No. 1.
- Ghomeshi, J. & Diane Massam (1994) Lexical/Syntactic Relations Without Projection. *Linguistic Analysis*, Vol. 24, No. 3-4.
- Ghomeshi, J. (1996) Projection and Inflection: A Study of Persian Phrase Structure. Ph. D. Dissertation, University of Toronto. Distributed by Toronto Working Papers in Linguistics.
- Goldberg, E. Adele (1995) Persian Complex Predicates and the Lexicon. Draft, University of California, San Diego.
- Grimshaw, J. & Mester, A. (1988) Light Verbs and Θ -Marking. LI 19, 205-232.

- Hale, K. & S. J. Keyser (1987) A View from the Middle. Lexicon Project Working Papers No. 10, Center for Cognitive science, MIT, Cambridge, Mass.
- Hale, K. & S. J. Keyser (1991) On the Syntax of Argument Structure. Lexicon Project Working Papers 34, MIT.
- Hale, K. & S. J. Keyser (1993) On Argument structure and the Lexical Expression of Syntactic Relations. In Hale, K. & J. Keyser (eds.) *The View from Building 20*. The MIT Press, Cambridge.
- Hale, K. & S. J. Keyser (1994) The Limits of Argument Structure. ms. MIT.
- Halle, M. & Marantz, A. (1993) Distributed Morphology. In Hale, K. & J. Keyser (eds.) *The View from Building 20*. The MIT Press, Cambridge.
- Heim, R. I. (1982) The Semantics of Definite and Indefinite Noun Phrases. Ph.D. dissertation, University of Massachusetts.
- Heny, J. & V. Samiian (1991) Three cases of Restructuring in Modern Persian. In K. Hunt, T. Perry & V. Samiian (eds.) *Proceedings of the Western Conference in Linguistics*.
- Higginbotham, J. (1985) On Semantics. LI 16.
- Higginbotham, J. (1987) Indefiniteness and Predication. In Eric J. Reuland & Alice G. B. ter Meulen (eds.) *The Representation of (In)definiteness*. The MIT Press.
- Hoop, de H. (1990) Restrictions on Existential Sentences and Object-Scrambling: Some Facts From Dutch. In Aaron L. Halpern (ed.) *The Proceedings of WCCFL 9*.
- Hoop, de H. (1992) Case Configuration and Noun Phrase Interpretation. Ph.D. dissertation, University of Groningen.
- Hopper, Paul J. & Thompson Sandra A. (1980) Transitivity in Grammar and Discourse. In *Language* 56, No. 2, 251-299.
- Jackendoff, R. (1990) *Semantic Structures*. MIT Press, Cambridge, Massachusetts.
- Jaeggli, A. O. (1986) Passive. *Linguistic Inquiry* 17, No. 4.
- Khanlari, P.N. (1979) *A history of the Persian Language*, vol. II. Translated by N.H. Ansari. Sterling Publishing Pvt Ltd, New Delhi.
- Larson, K. R. (1987) On the Double Object Construction. Lexicon Project Working Papers 16. Lexicon Project Center for Cognitive Science, MIT.
- Milsark, L. G. (1974) Existential Sentences in English. Ph.D. Dissertation, MIT. distributed by Indiana University Club.

- Milsark, L. G. (1977) Toward an Explanation of Certain Peculiarities of the existential Constructions in English. In *Linguistic Analysis* 3, No. 1.
- Mohammad, J. & S. Karimi (1992) 'Light' Verbs are Taking Over: Complex Verbs in Persian. In J. Nevis, V. Samiian, & Fresno (eds.) *Proceedings of the Western Conference on Linguistics*.
- Mohanan, T. (1995) Wordhood and Lexicality: Noun Incorporation in Hindi. *Natural Language & Linguistic Theory* 13, 75-134.
- Moyne, A. J. (1970) The Structure of Verbal Constructions in Persian. Ph.D. dissertation, Harvard University.
- Partee, B. (1987) Noun phrase interpretation and type shifting principles. In J. Groenendijk, D. de Jongh, and M. Stokhof (eds.) *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, Dordrecht: Foris.
- Rothstein, D.S. (1983) The Syntactic Forms of Predication. MIT Ph.D. dissertation. Reproduced by the Indiana University Linguistics Club (1985).
- Safir, J. K. (1985) *Syntactic Chain*. Cambridge University Press.
- Safir, J. K. (1987) What Explains the Definiteness Effect. In Eric J. Reuland & Alice G. B. ter Meulen (eds.) *The Representation of (In)definiteness*. The MIT Press.
- Shibatani, M. & Taro Kageyama (1988) Word Formation in a Modular Theory of Grammar. *Language* 64, No. 3.
- Siewierska, A. (1984) *The Passive: A Comparative Linguistic Analysis*. Croom Helm (Linguistic Series).
- Speas, M. (1990) *Phrase Structure in Natural Language*. Kluwer AC.
- Sproat, R. (1988) On Anaphoric Islandhood. In M. Hammond & M. Noonan (eds.) *Theoretical Morphology: Approaches in Modern Linguistics*. Academic Press, Inc. (1988).
- Sproat, R. (1985) On Deriving the Lexicon. Ph.D. thesis, MIT.
- Szabolcsi, A. (1984) From Definiteness Effect to Lexical Integrity. In Abraham, W. & S. d. Meij (eds.) *Topic, Focus, and Configurationality*. John Benjamin Publishing Company (1986).
- Szabolcsi, A. (1986) Indefinites in Complex Predicates. In *Theoretical Linguistic Research* 2.

- Tabaian, H. (1979) Persian Compound Verbs. *Lingua* 47, 189-208.
- Talmy, L. (1985) Lexicalization Patterns. In T. shopen (ed.) *Language Typology and Syntactic Description*, CUP. PP. 57-149.
- Tenny, L. C. (1987) Grammaticalizing Aspect and Affectedness. Ph.D. Thesis, MIT
- Vahedi, M.-M. (1996) The Syntax, Semantics, and Argument Structure of Complex Predicates in Modern Farsi. Ph.D. dissertation, University of Ottawa.
- Webelhuth, G. & Ackerman, F. (1992) Complex Predicates and Wordhood: passive Constructions in German. In CLS 28.
- Williams, E. (1994) Thematic Structure in Syntax. *Linguistic Inquiry Monograph* 23. The MIT Press.
- Windfuhr, L. G. (1979) *Persian Grammar: History and State of its Study*. Mouton Publishers, The Hague.