



A corpus study on identification and semantic classification of light verb constructions in Persian: the case of the light verb *xordan* 'to eat/collide'



Ramin Golshaie*

Iranian Research Institute for Information Science and Technology (IranDoc), Tehran, Iran

ARTICLE INFO

Article history:

Received 17 May 2015

Received in revised form 18 April 2016

Accepted 17 May 2016

Available online 11 June 2016

Keywords:

Light verb constructions

Corpus linguistics

Behavioral profile analysis

Collocations

Semantics

ABSTRACT

One of the challenging problems in the domain of Persian light verb constructions (LVCs) is to discover and classify a light verb (LV)'s multiple senses. This is important since productive use of LVs in Persian that leads to formation of novel LVCs can be explained in reference to these established LV senses. On the other hand, identification of LVCs in the first place is another problem which is a complex task given that not only no objective criterion exists for their identification but also the constituent elements of some LVCs can be split by interposing linguistic units that makes their identification difficult. This paper addresses these two issues using corpus methodology. To identify LVCs, the LV *xordan* with two unrelated meanings 'to eat/collide' was chosen for analysis and the corresponding LVCs were extracted from a sampled 50-million-word corpus based on a measure of collocational associations. The extracted LVCs consisted of frequent compositional and idiomatic noun-verb (N-V) patterns found in the corpus. Corpus examinations revealed that frequent compositional N-V sequences have constructional meanings and need to be recognized as LVCs. Finally, to discover the LV senses, 700 concordance lines of the extracted LVCs were studied and classified based on a behavioral profile analysis of their corpus usage patterns. According to the results of behavioral profile analysis, two constructional senses EAT and COLLIDE are coexistent under *xordan* each subsuming their own semantically-related LVCs. The findings while supporting the overall constructionist assumptions on polysemy network of LV senses necessitate a reconsideration of constructionhood criteria in Persian LVCs alongside the process of identification and classification of senses.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Persian light verbs constructions¹ (LVCs)² have been the focus of much controversial research in the recent decades. These multiword expressions are composed of a non-verbal element (NV), which is usually a noun or an adjective, and a light verb (LV) which is semantically impoverished compared to its 'heavy' counterpart. In studying Persian LVCs, various aspects

* P.O. Box: 13185-1371, 1090 Enqelab Ave., Tehran, Iran. Tel.: +98 (0)21 66 4949 80.

E-mail address: golshaie@irandoc.ac.ir.

¹ Also known as 'compound verbs' or 'complex predicates' in the literature.

² The following abbreviations have been used in the article: 1/2/3 = person marker; ACC = accusative marker; lit. = literal; LV = light verb; LVC = light verb construction; N = noun; NV = non-verbal element; PAST = past tense; PL = plural; POSS = possessive; PROG = progressive; SG = singular; V = verb. Other abbreviations have been defined in the text or footnote where necessary.