



Faculty of Human and Social Sciences

# **Persian Idioms Collection and Identification in Texts**

Vahid Ostovar

Dissertation  
Master project in **Language Sciences**

**Supervisor: Prof. Doutor Jorge Baptista**

**Faro, 2014**

# Statement

## *Persian Idioms Collection and Identification in Texts*

### **Declaração de autoria do trabalho**

Declaro ser o autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

©2014, Vahid Ostovar/ Universidade do Algarve

A Universidade do Algarve tem o direito, perpétuo e sem limites geográficos, de arquivar e publicitar este trabalho através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, de o divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

## *Persian Idioms Collection and Identification in Texts*

### **Declaration of authorship of the work**

I hereby declare to be the author of this work, which is original and unpublished. Authors and works consulted are properly cited in the text and listed in the list of references here included.

©2014 Vahid Ostovar / University of Algarve

The University of Algarve is entitled, perpetually and without any geographical limits, to file and publicizing this work through printed copies reproduced on paper or in digital form, or by any other known medium or any as yet to be invented, through its promotion on scientific reposition and admit its copy and distribution for educational or research non-commercial purposes, as long as credit is given to the author and to the publisher.

## **Dedication/Acknowledgments**

I would like to express my special appreciation and thanks to my advisor Professor Dr. Jorge Baptista, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a researcher. Your advice on both research as well as on my career have been priceless.

A special thanks to my family. Words cannot express how grateful I am with my mother, my father, my mother-in-law, and father-in-law for all of the sacrifices that you have made on my behalf. Your prayer for me was what sustained me thus far. I would also like to thank all of my friends who supported me in writing, and incited me to strive towards my goal. At the end, I would like to express appreciation to my beloved wife Bahareh Kakanaeni who spent sleepless nights with and was always my support in the moments when there was no one to answer my queries.

## **Abstract**

An idiom is a string of words whose meaning is different from the meaning conveyed by its individual words. This project will study Persian idioms with the structure *NO CI V*, which are sentences with a free subject (NO), a frozen direct object (CI) and a verb (V). The purpose of this project is to build a database of Persian idioms in order to be used in the computational processing of this language. First, a selection of web sources was used for the collection of idioms; second, from this a database of Persian idioms was built; third, a set of finite-state tools was used to intersect the database with reference graphs and build FSTs (transducers) for corpus exploring; fourth, these FSTs were then used to extract idiom candidates from a large subtitles Persian corpus; fifth, the resulting candidates lists was evaluated in order to: (a) estimate the scope of the database; (b) determine the precision of the task of identifying the idioms, using the finite-state tools; and (c) compare it with two association measures (t-test and chi-square). Results show chi-square to be an efficient association measures to retrieve idioms candidates; however, the finite state tools allow for a better precision. Attention should also give to the idioms' main verb; namely, full verbs tend to yield more precise result than more grammaticalized verbs such as support verbs. The database, in its current state, contains 364 verbal idioms form a single formal class.

## **Keywords**

Persian, verbal idioms, multiword expressions, Corpus Linguistics, statistical association measures.

## **Resumo**

Uma expressão idiomática é uma sequência de palavras cujo significado é diferente do significado veiculado pelas suas palavras quando usadas individualmente. Este projeto estudou as expressões idiomáticas em língua Persa com a estrutura N0 V C1, que são frases com um sujeito livre (N0), um objeto direto fixo (C1) e um verbo (V). O objetivo deste projeto é construir um banco de dados (ou léxico-gramática) de expressões idiomáticas persas, a fim de ser utilizado no processamento computacional desta língua. Em primeiro lugar, uma seleção de fontes da web foi utilizada para a recolha de expressões idiomáticas; a partir desta recolha, procedeu-se à construção de um banco de dados de expressões idiomáticas persas; seguidamente, usaram-se ferramentas próprias para construir máquinas de estados-finitos de referência e intercetá-las com o banco de dados e assim produzir FSTs (transdutores ) capazes de explorar corpora textuais; estes FSTs foram então usados para extrair sequências candidatas ao estatuto de expressão idiomática partir de um corpus Persa de legendas de filmes; em quinto lugar, as listas de candidatos resultantes foram avaliadas com o fim de: (a) estimar a extensão da base de dados; (b) determinar a precisão da tarefa de identificação das expressões idiomáticas quando são usadas as ferramentas de estados finitos aqui empregues; e (c) comparar estes resultados com os que podem ser obtidos usando duas medidas de associação (o Student t-teste e o qui-quadrado). Os resultados mostram que o qui-quadrado é uma medida de associação eficiente para recuperar sequências candidatas a expressões idiomáticas; no entanto, as máquinas de estados finitos permitem uma melhor precisão; Também se deve dar atenção ao verbo principal envolvido nessas expressões idiomáticas, ou seja, verbos plenos tendem a produzir resultados mais precisos do que verbos mais gramaticalizados, como os verbos-suporte. O banco de dados, no seu estado atual, contém 364 expressões idiomáticas verbais, de uma única classe formal.

## **Palavras-chave**

língua Persa , Expressões idiomáticas verbais , expressões multipalavra , Linguística de Corpus, medidas de associação estatísticas.

## Extended Abstract

An idiom is a string of words whose meaning is different from the meaning conveyed by the individual words. This project studied Persian idioms with the structure N0 C1 V, which are sentences with a free subject (N0), a frozen direct object (C1) and a verb (V). The idioms were collected from Persian idiomatic expressions dictionaries and some websites. In order to do get those idioms from the websites, the search was limited to those websites that contains more than 100 Persian idioms each. The purpose of this project is to build a database (or lexicon-grammar) of Persian idioms in order to be used in the computational processing of this language. To reach the goal of this project the following steps were taken:

(1) Web source selection: In order to select the web sources for collecting Persian idioms the web has been queried, using Google and other browsers that are powered by Google, such as AVG. Moreover, the result and analysis systematically limited to the first top ten pages per query.

(2) Building a database of Persian idioms: The database of Persian idioms is an Excel™ file that contains 364 (Verb+C1) Persian idiomatic expressions. For each idiom, a transliteration, a gloss, a rough translation and exact meaning has been provided as well as the sources from which it was retrieved.

(3) Building FST tools for corpus exploring: A first evaluation of the database has been done by intersecting the tabular data with a finite-state reference transducer (FST), using the Unitex (3.0) corpus-processing tool; this process yields a set of FSTs that are then able to retrieve candidate expressions from texts.

(4) Corpus analyzing and retrieving idioms' candidates: A corpus of movie subtitles in Persian, which contains 556,234 sentences, 15,166,987 (64,492 diff.) tokens, and 4,485,147 (64,365) simple forms was used in this project. Lists of idioms' candidates were retrieved by using the FSTs mentioned in (3).

(5) Evaluation: to evaluate the results, three tasks have been done:

(i) Estimating the scope of the database: to estimate the scope of the database idioms, we produced a dummy FST that would extract all the instances containing the verb and the lexical constant C1 from the text, allowing for a window of length 0 to 5 words, which inserted between the verb and the constant. By intersecting the database with this FST, we were able to produce a set of transducers that extract all instances that “potentially” include the idioms of the database;

(ii) Determining the precision of the task of identifying the idioms, using the Unitex tools: precision means to try and measure how many of the idiomatic expressions retrieved from the corpus were correctly captured.

(iii) Association measures: for the idioms that were captured by the FST using the entire lexicon-grammar, two association measures were calculated: t-test and chi-square.

An evaluation of the results was produced, in terms of the FSTs precision, and by using the association measures; the two different approaches were also compared.

Results show that depending on the verb, namely, if the verb has a more grammatical status or is a full verb, it may be necessary to adjust insertion windows between the key elements of the idiom; precision also varies depending on the verb. Association measures are valid strategies to capture these types of idioms; however, their precision depends on the idioms’ verb, as the presence of support verbs may influence the results.

## Resumo Alargado

Uma expressão idiomática é uma sequência de palavras cujo significado é diferente do significado veiculado pelos seus quando usados individualmente. Este projeto estudou expressões idiomáticas em língua Persa, com a estrutura N0 V C1, isto é frases com um sujeito livre ( N0), um objeto direto fixo (C1) e um verbo (V). As expressões idiomáticas foram coligidas a partir de dicionários expressões idiomáticas e de alguns sites especializados; neste último caso, a busca foi limitada aos sites que contém mais de 100 idiomas persas cada. O objetivo deste projeto é construir um banco de dados (ou léxico-gramática) de expressões idiomáticas persas, a fim de ser utilizado no processamento computacional desta língua. Para alcançar o objetivo deste projeto as seguintes etapas foram percorridas:

(1) seleção de fonte Web: Para selecionar as fontes da web de onde seriam recolhidas as expressões idiomáticas persas, usou-se o Google e outros navegadores que são suportados pelo Google, tal como o AVG. Além disso, procedeu-se à análise sistemática dos resultados, limitando-os às primeiras dez páginas por consulta.

(2) Construir uma base de dados de expressões idiomáticas persas : O banco de dados de expressões idiomáticas persas é um arquivo Excel™ que contém 364 (verbo + C1) expressões. Para cada expressão, uma transliteração, uma glosa, uma tradução aproximada e o significado exato foram fornecidos, bem como as fontes de onde esta foi recolhida.

(3) ferramentas: Construção de FST para exploração de corpora: Uma primeira avaliação do banco de dados foi feita através da interseção dos dados em formato tabular com um transdutor de estados finito (FST) de referência, usando a ferramenta de processamento de corpus Unitex (3.0); este processo produz um conjunto de FSTs que são então capazes de recuperar cadeias candidatas ao estatuto de expressões idiomáticas a partir de textos.

(4) Análise do corpus e recuperação de cadeias candidatas a expressões idiomáticas : um corpus de legendas de filmes em Persa, que contém 556.234 frases, 15.166.987 (64.492 diferentes) tokens, e 4.485.147 (64.365 diferentes) formas simples foi utilizada neste projeto. As listas de candidatos a expressões idiomáticas foram recuperadas usando os FSTs mencionados em (3).

(5) Avaliação : para avaliar os resultados, três tarefas foram executadas:

(i) Estimar o escopo do banco de dados: para estimar o alcance das expressões idiomáticas de banco de dados, produzimos um FST de teste que extrai do texto todas as instâncias que contêm o verbo e a constante lexical C1, permitindo uma janela de comprimento de 0 a 5 palavras entre o verbo e essa constante. Através da interseção do banco de dados com esse FST de teste, fomos capazes de produzir um conjunto de transdutores que extraem todas as instâncias que "potencialmente" incluem as expressões idiomáticas do banco de dados;

(ii) Determinar a precisão da tarefa de identificação das expressões idiomáticas, utilizando as ferramentas do sistema Unitex : precisão significa tentar medir quantas das expressões idiomáticas recuperados do corpus foram corretamente capturadas.

(iii) Medidas de associação : para as expressões idiomáticas que foram capturadas pelo FST usando o léxico-gramática inteiro, foram calculadas duas medidas de associação; o Student t-teste e o chi-quadrado .

A avaliação dos resultados foi produzida em termos de precisão dos FSTs e usando as medidas de associação; as duas diferentes abordagens foram também comparadas.

Os resultados mostram que, dependendo do verbo, ou seja, se o verbo tem um estatuto mais gramatical ou é um verbo pleno, pode ser necessário ajustar as janelas de inserção entre os elementos-chave da expressão; a precisão também varia dependendo do verbo. Medidas de associação são estratégias válidas para capturar esses tipos de expressões idiomáticas, sobretudo o chi-quadrado; no entanto, sua precisão depende verbo das expressões idiomáticas, já que a presença de verbos-suporte pode influenciar os resultados.

# Contents

1. Introduction.....	1
1.1. Aims and Objectives.....	2
2. Theoretical Background.....	3
2.1. What is an idiom? .....	3
2.1.1. The Structure of Idioms .....	5
2.1.2. Types of Idioms .....	6
2.1.3. Characteristics of Idioms .....	6
2.2. What is collocation? .....	7
2.2.1. Structural types of collocation .....	8
2.2.2. Types of collocation .....	9
3. Methodology .....	10
3.1. Selection of web sources .....	10
3.2. Building a database of Persian idioms .....	13
3.3. Building FST tools for corpus exploring .....	14
3.3.1. Reference graph.....	14
3.3.2. Intersecting the reference graph with the database.....	17
3.4. Corpus parsing and idioms candidate extraction .....	21
4. Evaluation .....	22
4.1. Estimation of idioms database scope.....	22
4.2. Determine precision of identification by using the Unitex tools .....	29
4.3. Association measures.....	30
5. Conclusion .....	32
References.....	34
Active bibliography .....	34
Dictionaries.....	35
Sources .....	36
Other sources .....	36
Appendices .....	38
Appendix A. Database.....	38
Appendix B. Dictionary of the inflected forms (کرد, <i>kard</i> , کن, <i>kon</i> “to do”).....	62
Appendix C. Dictionary of the inflected forms (زد, <i>zad</i> , زن, <i>zan</i> “to hit”).....	63

## List of tables

Table 1. Different Definitions of Idiom .....	3
Table 2. Different Definitions of collocation.....	7
Table 3. Database of Persian idioms (extract) .....	13
Table 4. Database with the past and present stem of each idiom verb.....	22
Table 5. Experiments with the idioms involving verbs کردن <i>kardan</i> (to do) and زدن <i>zadan</i> (to hit).....	26
Table 6. Concordances result of each dummy FSTs.....	27
Table 7. Precision.....	29
Table 8.” کردن, <i>kardan</i> (to do)”T-test and chi-square .....	30
Table 9.” زدن <i>zadan</i> (to hit).”T-test and chi-square .....	31
Table 10. Comparison of t-test and chi-square .....	31

## List of figures

Figure 1. Reference graph: C1ref.grf.....	15
Figure 2. Unicode database.....	18
Figure 3. C1res.grf .....	19
Figure 4. C1res.0179.grf .....	19
Figure 5. C1res.0187.grf .....	20
Figure 6. Concordance .....	21
Figure 7. Extract.ref.grf .....	23
Figure 8. Concordance of dummy FST.....	23
Figure 9. Dummy reference FST for idioms with the verb “ کردن <i>kardan</i> (to do) ” with 0-5 insertion window.....	24
Figure 10. Reference FST for idioms with the verb “ کردن <i>kardan</i> (to do) ” without any insertion .....	25
Figure 11. Dummy reference FST for idioms with the verb “ زدن <i>zadan</i> (to hit), ” with 0-5 insertion window.....	25
Figure 12. Reference FST for idioms with the verb “ زدن <i>zadan</i> (to hit), ” without any insertion .....	26

## 1. Introduction

An idiom is a multi-word expression, according to Larson (1984, p. 20) an idiom is a string of words whose meaning is different from the meaning conveyed by the individual words. While, a collocation is a multi-word expression, that are words or a phrase which is often utilized with another word or phrase, in a way that sounds correct to users of the language but, the frequency of that Co-occurrence is higher than coincidentally Co-occurrence. Some researchers consider idioms as a collocation, Lewis (2000, p. 51) believes that It would not be erroneous if we claim that all collocations are idiomatic and all phrasal verbs and idioms are collocations or contain collocations, as all are expectable combinations of different kinds.

A collocation is a special, non-trivial way that some words combine in a language to produce natural speech and writing. For example, when you say “*pay attention*” it could be said as “*give attention*”, or “*put attention*”, but it is not: it is “*to pay attention*” because; the verb “*to pay*” has a special status with the noun “*attention*”, which can be determined by the fact that, this particular combination occurs with a frequency higher than were chance. However, an idiom is a sequence of words whose meaning is different from the meaning of the each single word when used independently. For example, when you say, “*jack hit the ceiling*” you do not truly denote that jack jumped and hit the ceiling with his hand or foot but that means “*jack really gets angry*”.

Many languages in the world have idioms. Like many other languages, Persian is a language that has those types of expressions (idioms) as well and several different constructions can be found. It is important to determine accurately what kind of expressions is considered in this project. This project will study Persian idioms with the structure *NO CI V*, which are sentences with a free subject (N0), a frozen direct object (C1) and a verb (V). For instance:

Persian:	وحدید	تاب	خورد
Transliteration:	Vahid	tab	khord-past
Translation:	Vahid	swing	eats
Gloss:	Vahid	had	fun

This example has a *NO CI V* structure, (وحدید, *Vahid* is the subject N0, تاب *tab* (*swing*) is the direct object C1 and خورد, *khord* (*eats*) is the verb V). In this project, all idioms will be collected from

dictionaries of Persian idiomatic expressions; in addition, this approach is complemented by perusing extant lists of idioms that are available in the internet and dictionaries of idioms in Persian.

### **1.1. Aims and Objectives**

The main goal or the aim of this dissertation project is to prepare a database of Persian idioms to be used in the computational processing of that language. In order to reach the aim of this project; we must complete these following steps:

- 1) Selection of web sources for the collection of idioms
- 2) Building a database of Persian idioms
- 3) Building FST tools for corpus exploring
- 4) Corpus parsing and idioms candidate extraction
- 5) Evaluation

## 2. Theoretical Background

### 2.1. What is an idiom?

An idiom retrieved from the Latin word “idioma<sup>1</sup>” which meaning is “special feature, special phrasing, special property and one’s own “. Different linguists used different terms to label the concept of idioms; For example: Howarth (1998) prompts for the term phraseology, while McCarthy (1998) chooses the traditional term idiom, Glaser (1984) used the term phraseological unit, Carter (1998) preferred the term fixed expressions, Moon (1997) picks up the term multi-word items and Larson (1984) calls the idioms as special collocations, or fixed combination of words.

The *Longman Dictionary of Language Teaching and Applied Linguistics* defines an idiom as “an expression which functions as a single unit and whose meaning cannot be worked out from its separate parts” (1992, p. 198). On the other hand, Jonathan Crowther in the *Oxford Advanced Learner’s Dictionary of Current English*, expresses idiom as “a phrase or sentence whose meaning is not clear from the meaning of its individual words and which must be learnt as a whole unit.” (1995, p. 67)

Other scholars’ present dissimilar views of the idiomaticity notion, these views do not obligatorily disprove each other; rather, they may complement each other to engender a broad perspective on idioms as in the following table we present some of these definitions.

**Table 1. Different Definitions of Idiom**

<b>Author</b>	<b>Definition</b>
Crystal, David. (1985)	An idiom is a term utilized in grammar and lexicography to refer to a string of words that is semantically and often syntactically limited, so that they act as a single unit.
McMordie. W (1983, p. 4)	An idiom is a number of words, when they are together, denote something different from the each single word of the idiom when they stand alone.
Dean Curry. (1994, p. 49)	An idiom is the assigning of a fresh meaning to a set of words that already have their own meaning.
Moon. (1998, p. 3)	An ambiguous term, utilized in different ways. In general,

<sup>1</sup> <http://en.wikipedia.org/wiki/Idiom>

<b>Author</b>	<b>Definition</b>
	use, the idiom has two main denotations. First, an idiom is a specific denotes of expressing something in language, music, art, and so on, which describes a person or group. Second, an idiom is a specific lexical collocation or phrasal lexeme, unusual to a language
Collins. (2000, p. 70 )	An idiom is a set of words that is identified as a single unit, and has a different meaning from the designation of its single words and they customarily employed in everyday language to exactly express ideas and concepts that cannot be compressed into a single word.
Baker. (1992, p. 63)	Frozen forms of language that permit little or no variation in form and often carry meanings, which cannot be realized from their individual components.
Hatch and Brown. (1995, p. 202)	The set end of the continuum from loose to set collocation.
Larson. (1984, p. 20)	An idiom is a string of words whose meaning is different from the meaning conveyed by the individual words
Carter. (1987, p. 66)	Defined an idiom as follows: a) They are non-replaceable collocations. b) They usually occur as multiword components. c) They show semantic opacity.
Wright (1999, p. 7)	An idiom is an expression with double features. First, the idiom is fixed and is recognized by native speakers. You cannot make up your own. Second, It uses language in a non- literal- metaphorical way.
Benson, M., Benson, E. and Ilson, R. (1986a)	Idioms are relatively frozen expressions whose meanings do not reflect the meanings of their component parts
Cruse (1986)	To him an idiom is a lexically involute unit composed of a single minimal semantic constituent
Aisenstadt (1979)	An idiom is described as one semantic unit.
Oxford Dictionary (1983)	It distinct idiom in two parts: “pure idioms” (totally fixed) and “figurative idioms” (allowing for some variation)
Nattinger & DeCarrico, (1992)	He defined idiom as complex bits of frozen syntax whose meaning cannot be taken from the meaning of their constituents, that is, whose meanings are more than simply the totality of their single components.

As it can be seen, there is not necessarily any conflict between the above mentioned definitions of idiom; i.e. regardless of conspicuous differences in the way of expressing the concepts, they intersect at one point: an idiom is a fixed expression whose meaning cannot be worked out by merging the literal meaning of its individual words.

### 2.1.1. The Structure of Idioms

Mona Baker (1992, p. 63) describes five formal changes that customarily cannot be done to an idiom, they are as follows:

- 1) Changing the idioms word order. For example the idiom “*let the cat out of the bag*” that means you let someone know a secret, cannot be reordered to “*the cat got left out of the bag*”
- 2) Delete a word from the idiom. “*Let the cat of the bag*”
- 3) Add a word to the idiom. “*Let the cat and dog out of the bag*”
- 4) Replace one of the idiom words by another. “*Let the cat out of the briefcase*”
- 5) Changing the idioms grammatical structure. “*Let the cat out of a bag*”

According to McMordie (1983, p. 89), idioms divide into three different groups, as follows:

1. Irregular structure, but the meaning is clear and foreseeable from the words used (e.g. *make someone proud*).
2. Regular structure, but the meaning cannot be retrieved from the words used (e.g. *bring the house down*), which means to cause an audience to leap to its feet
3. Irregular structure, and the meaning cannot be realized from the individual words (e.g. *go great guns*), which means to be very successful.

Most idioms are unique and fixed in their grammatical structure. They could be adjectival, adverbial, verbal, or nominal. Fernando (1994, p. 223) defines six different types of combination that are: 1) pairs of adjectives, e.g. *high and dry*. 2) Pairs of nouns, e.g. *life and limb*. 3) Compound adjectives, e.g. *a freestanding tower, tongue-tied child*. 4) Collective noun and phrases, e.g. *A zoo of wild animals, an army of soldiers*. 5) Adjective noun phrases, e.g. *accomplished task, young mind*. 6) Noun phrases, e.g. *a bill of fare, a flag of truce*.

### 2.1.2. Types of Idioms

According to Fernando (1994, p. 35) he distinguishes three sub-classes of idioms:

1) Pure Idioms: Fernando defines a pure idiom as ‘a type of conventionalized, non-literal multiword expression’ (id, 36) Pure idioms are non-literal, however they might be either invariable or have little difference. Moreover, idioms are verbally expressed to be obscure (id, 32). For example, the pure idiom to *spill the beans* has nothing to do, either with *beans* or with *spill*.

2) Semi-Idioms: Semi-idioms are verbally expressed to have one or more literal constituent and one with non-literal sub-sense. Hence, this category of idioms is considered partly non-transparent (id, 60). For example, *foot the bill*, which means ‘pay’ is a semi-idiom. Because the word “*bill*”, which is an object in this example, has its own literal meaning, and only the word *foot* is idiomatic.

3) Literal Idioms: This sub-class of idioms are either invariable or permit little variation. Furthermore, literal idioms are considered to be crystal clear as they can be interpreted on the base of their components. For example: *in any case, on the other hand, in other words*.

### 2.1.3. Characteristics of Idioms

For idioms, linguists have suggested different features. According to Stock (1989, p.9), “on two fundamental perspectives: one perspective considers idioms as the rudimental units of language, with holistic characteristics, and the other perspective, which accentuates on the definite role in the consummate idiom”.

#### 1. Lexical and syntactic flexibility / Syntactic modifications

“A constrained number of idioms have one (or more) lexical variants. However, most are lexically frozen (non-productive) to a sizably voluminous extent. Analyzable (less idiomatic) expressions are often more open to lexical supersession and syntactic variation” (Fazly and Stevenson, 2006, p. 2).

Chang & Fischer (2006, p. 1) believe that “Many idioms are lexically specific and do not permit replacement of even close synonyms”. However, they further add, “that even idioms with fixed lexical items frequently appear in syntactically different forms” (id, p. 1).

## 2. Transparency / Non-compositionality / Opacity

“Transparency or semantic analyzability refers to the degree of agreement between the literal and figurative meanings of an idiom” (Cain, Oak Hill and Lemmon, 2004, p. 3).

“The meaning of an idiom is normally not foreseeable from those of its components, as in the often mentioned example *kick the bucket* (die). In contrast, *spill the beans* is decomposable into its components with *beans* modeling a secret and *spill* is revelation” (Chang & Fischer, 2006, p. 1).

“Idioms such as *by and large* and *spick and span* are often considered opaque: the semantic motivation or source of the idiom may be obscure, and no particular underlying image may come to mind” (id, p. 1). Chang & Fischer (id, p. 1) finally, conclude, “In their central (or prototypical) case, idioms are non-compositional, lexically simple.

## 3. Idiom familiarity

It has been defined by Nippold and Taylor (1995, p. 427, as quoted in Cain, Oak Hill and Lemmon, 2004, p. 2) as “how often an idiom happens in the language. Children, youths, and adults realized that more familiar idioms are easier to understand than less familiar ones”.

It might be deduced that: as, an idiom is a fixed expression whose meaning cannot be taken as a sum of the meanings of its component parts, idioms are non-literal expressions but figurative ones

## 2.2. What is collocation?

Collocation is the association between two words or sets of words that frequently go together and make a widespread expression (Williams, (2003)).

In the following table, some researchers’ interpretations about the notion of collocation are presented.

Table 2. Different Definitions of collocation

Author	Definition
Benson, M., Benson, E. and Ilson, R. (1990).	A collocation is an arbitrary and recurrent word combination

Cruse (1986)	Collocation is a sequence of lexical items which usually co-occur
Aisenstadt (1979)	A collocation is a combination whose constituents is utilized in one of their conventional, non-idiomatic meaning and track certain structural designs.
Oxford Dictionary (1983)	The “non-idioms” (that is, collocations) are divided into “restricted” (or “semi-idioms”), which allow a degree of lexical variation and in which some elements are used figuratively, and “open”, in which elements are freely combined and are used in a common literal sense
Nattinger & DeCarrico, 1992)	Collocations themselves are presented as “roughly predictable, yet restricted to certain specified items”.

To sum up, the following examples show the differences between researchers’ theories about the concepts of idiom and collocation:

- 1) Aisenstadt (1979) and Cruse (1986), look up to collocations as non-idioms. Benson, Benson, and IlsonBenson, M. (1990), regard collocations and idioms as shading in to each other. Nattinger & DeCarrico, (1992, p.20, 181), see idioms as a sub-group within collocations.
- 2) Nattinger & DeCarrico, (1992, p. 178), present “*kick the bucket*” as an example of collocation, while, the Oxford Dictionary (1983) and Cruse (1986, p. 37) presents the same example as a typical idiom.

### 2.2.1. Structural types of collocation

According to Benson, Benson, and Ilson (1986, p. ix), typical lexical collocations consist of nouns, adjectives, verbs, and adverbs, which are classified lexical collocations by structural types. In the following examples, the base (literal) word is in bold, and the word or expression that collocates with it is in italics:

(1) Verb + noun: wind a *watch*, quench one’s *thirst*

(2) Adjective + noun: torrential *rain*

Persian idioms

(3) Noun + verb: volcanoes *erupt*

(4) Noun1 + of + noun2: a school of *whales*

(5) Adverb + adjective: closely *acquainted*

(6) Verb + adverb: *apologize* sincerely

### 2.2.2. Types of collocation

According to Hill, J. (2000, p. 63, 64), collocations categorized into four groups in view of the strength of the association between the words:

- 1) Weak Collocations: They occur between two common words, each of which may co-occur with many other words. e.g. “*white shirt*” or “*expensive watch*”
- 2) Medium strength collocations: They are the ones, which are utilized with prevalent words like *make, do, keep*, etc.; e.g. “*Make a mistake*”, “*do someone a favor*”
- 3) Strong collocations: the collocations, which are regularly used together; e.g. “*trenchant criticism*”.
- 4) Unique collocations: Words which have no other potential collocation; e.g. “*foot the bill*”

The strength of the collocation can be quantified by using association measures, (Manning Chr. and Schütze H, 2005, chapter, 5) such as mutual information or through hypothesis testing like student t.test and chi-square. Pecina 2009 made an overview of the most commonly used association measures for collocation extraction.

### 3. Methodology

As mentioned above, all idioms for the study are collected from Persian idiomatic expressions dictionaries such as (*Persian Slang Dictionary*) and some websites, In order to do this; we limit our search to those websites that contains more than 100 Persian idioms each. Then we built a collection of lexicon entries in the database, which are the idioms with *NO CI V* structure. Afterwards, we preside with the formalization of the distributional, structural and transformational properties of the idioms.

Finally, with this database, we will take initial steps towards the automatic identification of these idioms in texts; our goal is to determine the main issues in this process. By intersecting these data with FST, a preliminary assessment of the lexical database, used in capturing idioms appearing in Persian texts, will be carried out.

#### 3.1. Selection of web sources

In order to select the web sources for collecting Persian idioms, we queried the webs, using Google and other browsers that are powered by Google, such as AVG<sup>2</sup>, from January 2013 to May 2013. Moreover, we have systematically limited our result and our analysis to the first top ten pages per query. These are the most relevant queries that have been used.

Q1) [فرهنگ اصطلاحات عامیانه زبان فارسی]

Farhange estelahate amiane zabane Farsi

Dictionary terms slangy language Farsi

“Persian language expressions dictionary”

Q2) [اصطلاحات عامیانه و رایج به زبان فارسی]

Estelahate amiane va rayej be zabane Farsi

Terms slangy and common in language Farsi

---

<sup>2</sup> <http://www.avg.com/eu-en/secure-search>

“Common expressions in Persian language”

Q3) [فرهنگ اصطلاحات زبان فارسی]

Farhange estelahate zabane Farsi

Dictionary terms language Farsi

“Expressions dictionary of Persian language”

We have found several sites by going systematically through the results from these queries above, although, several hits were discarded as not relevant, namely:

1) Sites about English idioms and other languages in Farsi for example:

Website name: Expressions and interpretations

URL: <http://www.englishpro.ir>

Date Accessed: February 25, 2013

2) Sites that did not contain enough idiomatic expressions and are really poor to be considered as a source, for example:

Website name: Persian slang terms and their meanings

URL: <http://www.noandishaan.com/forums/thread40805.html>

Date Accessed: February 25, 2013

3) Many of them just give some general information about Persian idiomatic expressions and advertise dictionaries of idioms that will be published in the future, for example:

Website name: jungle publication.

URL: <http://www.junglepub.org>

Date Accessed: February 25, 2013

- 4) Websites that contain old Persian expressions.

Website name: Expression glossary (in Qajar period).

URL: <http://www.persianacademy.ir>

Date Accessed: February 25, 2013

- 5) Sites appeared more than once in each query.

All sites were mentioned appeared in references.

We retained the following relevant websites as a valid source, because they contain sufficient (more than 100) idiomatic expressions.

- Adib, A. (2010). *Persian Language and Literature* [zaban va adabiate Farsi]. Retrieved February 19, 2013, from <http://aryaadib.blogfa.com/cat-30.aspx>
- Samaee, Seyed Mehdi. (2003) [farhange estelahate zabane makhfi] فرهنگ اصطلاحات زبان مخفی *Hidden Language Glossary*. (D. A. Sadeghi, Ed.) (First., p. 101). Tehran: Ghazal. Retrieved from <http://forum.p30world.com/showthread.php?t=559720>
- S/N. (2013). [farhange estelahate amiane] فرهنگ اصطلاحات عامیانه *Expressions Dictionary*. Retrieved February 19, 2013, from <http://www.sootak.ir>
- S/N. (2013). [estelahate pamiane va rayej be zaban farsi va engelisi] اصطلاحات عامیانه و رایج به زبان فارسی و انگلیسی *Popular Idiomatic Expression in Persian and English*. Retrieved February 19, 2013, from <http://hedionline.com/index.php/learn-english/phrasal-verb-idiom-slang/142-2012-11-10-19-45-33>

- Fararoy, J. (2009). [farhange Teyfi] فرهنگ طیفی *Teyfi Dictionary* (p. 747). [nashre Hermes] نشر هرمس Hermes publication. Retrieved from <http://www.scict.ir/portal/Home/Default.aspx?CategoryID=a88bdb56-4ebd-4c32-80d0-a8406cd1655d>. Date Accessed:23 February 2013

### 3.2. Building a database of Persian idioms

The database of Persian idioms is an Excel™ file that contains 364 Persian idiomatic expressions. A small number of them are represented in Table 3 and the others are observable in the appendix. The database took the shape of the table, where lines represent the idiomatic expressions and the column its linguistic properties. For each idiom, a transliteration, a gloss, a rough translation and exact meaning has been provided as well as the sources from which it was retrieved. The linguistic properties were consisting of: a) the distributional constraints on the subject, the possibility of passive. b) The idiom itself was encoded; one element per column, and special attention was given to the variation of determiners and modifiers, as well as other relevant features. c) The probability of being positive or negative literal, it means that, whether each idiom has a literal meaning beside of its own idiomatic meaning or not.

Table 3. Database of Persian idioms (extract)

NO:Hum	NO:Nhum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	آب	<E>	آوردن	<E>	-	+	آب آوردن	ab avaridan	water bring	bringing water	being sick	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	آب	<E>	دادن	<E>	-	+	آب پس دادن	ab pas dadan	water give back	giving back the water	being generous -being unreliable	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	+	آب	<E>	رفتن	<E>	-	+	آب رفتن	ab raftan	water go	going water	became short-became low	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	آتش	<E>	کردن	<E>	-	+	آتش کردن	atash kardan	fire do	making fire	shooting	<a href="http://www.scict.ir">http://www.scict.ir</a>

NO=:Hum	NO=:Nhum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	آتش	<E>	گرفتن	<E>	-	+	آتش گرفتن	atash gereftan	fire take	taking fire	get angry	<a href="http://www.scict.ir">http://www.scict.ir</a>

### 3.3. Building FST tools for corpus exploring

A first assessment of the database is done by intersecting the tabular data with Finite state transducer (**FST**), using the Unitex (3.0) corpus-processing tool. “**Unitex**<sup>3</sup> is a corpus processing system, based on automata-oriented technology, with this tool; you can handle electronic resources such as electronic dictionaries and grammars and apply them. You can work at the levels of morphology, the lexicon and syntax. The main functions of Unitex are: building, checking and applying electronic dictionaries, pattern matching with regular expressions and recursive transition networks, applying lexicon-grammar tables, handling ambiguity via the text automaton, aligning text, building an automaton from corpus”. By **Corpus**<sup>4</sup>, we mean “a collection of linguistic data, either compiled as written texts or as a transcription of recorded speech. The main purpose of a corpus is to verify a hypothesis about language for example, to determine how the usage of a particular sound, word, or syntactic construction varies”.

#### 3.3.1. Reference graph

In this part, at first we have built our main (reference) graph, which is shown in Figure 1. This graph reads from right to left.

<sup>3</sup> <http://www-igm.univ-mlv.fr/~unitex/index.php?page=0>

<sup>4</sup> [http://www.tuchemnitz.de/phil/english/chairs/linguist/independent/kursmaterialien/language\\_computers/whatis.htm](http://www.tuchemnitz.de/phil/english/chairs/linguist/independent/kursmaterialien/language_computers/whatis.htm)

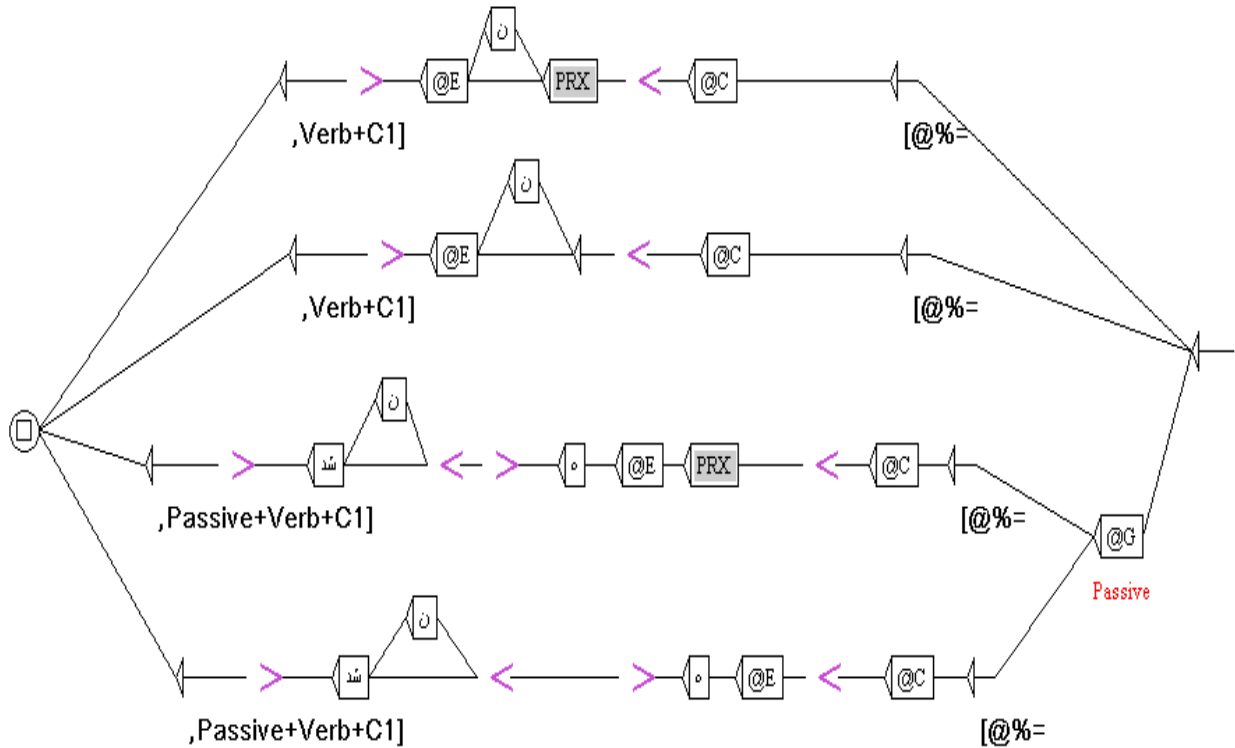


Figure 1. Reference graph: C1ref.grf

The variables introduced by the symbol “@” correspond to the columns of the Table 3, where, @C=Noun (direct object) and @E=Verb. The sub-graph PRX (in the gray boxes) includes the prefixes that are used to form some compound tenses of Persian verbs. These prefixes can be both separated from the verb as autonomous simple words, or joined to the verb form. In this way the reference graph of Figure 1 shows this sub-graph in two paths; on the upper path it is separate from the node indicating the verb; on the second path it is presented in morphological mood, delimited by '<' and '>', so it represents the case where the prefix is joined with the verb form<sup>5</sup>.

The two lower paths are initiated by the node with @G, which corresponds to the property Passive in the lexicon-grammar table. In this table this is a binary property, indicated by the symbols ‘+’ and ‘-’ depending on whether the idiom accepts the Passive transformation or not, respectively, while keeping its idiomatic value. For example:

<sup>5</sup> At this time, we are not concerned with the inflection of verbs, which is the topic of another dissertation (Bahareh Kakanaeeni, *in preparation*). Therefore, the graphs was only worked with the infinitive forms.

1a) Persian: سنگ انداخت: وحید (+' passive idiom, Figure 4)

Transliteration: *Vahid sang andakht*

Translation: *Vahid is throwing stone*

Gloss: *Vahid is a trouble maker*

1b) ✓ Passive: *stone was thrown by Vahid (Vahid is a trouble maker)*

✓ سنگ توسط وحید انداخته شد

2a) Persian: وحید دم در آورد (-' passive idiom, Figure 5)

Transliteration: *Vahid dom daravard*

Translation: *Vahid evolves tail*

Gloss: *Vahid becomes shameless*

2b)\* passive: *tail is grown by Vahid*

\* دم توسط وحید درآ ورده شد

In this way the variable @G functions have a switch; if the idiom has the passive property with the '+' sign, the remainder of the path is built; otherwise, the graph collapses at that point. This allows us to precisely define which sentence structures are to be matched for each idiom. Furthermore, Persian passive sentences also involve a verbal suffix 'ه' (*he*) and the auxiliary verb 'شد' (*shod*). Besides this, we also added a prefix 'ن' (*ne*) to verb form in order to get the negative verbs.

In the output of this FST, we find two delimiters, [and], and a variable "@%", which will indicate the number of the line in the lexicon-grammar table of the idiom being processed. At the end of the output, the tags "[Verb + C1]" and "[Passive+Verb+C1]" are also added in this particular order as, the graph applied from right to left, while the correct order is "[Verb+C1," and "[Passive+Verb+C1,".

### **3.3.2. Intersecting the reference graph with the database**

Our idiom database has been saved as a Unicode comma-separated values text (Figure 2). We intersect this table with the reference graph and produce as output a resulting graph containing one sub-graph per idiom. The system automatically builds these graphs by systematically exploring each line of the table and using the reference graph, it replaces the variables by the content in the table. The system also numbers each sub-graph with the line number corresponding to the idiom in the table, Figure 3 with the label, result graph. In addition, two different sub-graphs have been shown in Figure 4 and Figure 5.

Persian idioms

N...	N...	C1	Det	Verb	Modify	Pa...	Lite...	Persian idioms	Transliteration	Gloss	Rough translation	Exact mean...	Reference
+	-	درز	<E>	گرفت	<E>	-	+	درز گرفتن	darz gerft	seam take	taking seam	talkin...	Persia...
+	-	دخوری	<E>	داشت	<E>	-	-	دخوری داشتن	delkhorī dasht	Annoyance have	having ann...	being ...	Persia...
+	-	دل دل	<E>	کردن	<E>	-	-	دل دل کردن	del del kard	heart heart do	doing hear...	Vacillate	Persia...
+	-	دم	<E>	آوردن	<E>	-	+	دم در آوردن	dom daravard	tail evolve	evolve tail	becomi...	Persia...
+	-	ذوق	<E>	کردن	<E>	-	-	ذوق کردن	zogh kard	taste do	doing taste	becomi...	Persia...
+	-	رکاب	<E>	دادن	<E>	-	+	رکاب دادن	rekab dad	pedal give	giving pedal	give up	Persia...
+	-	کف	<E>	رفتن	<E>	-	-	کف رفتن	kaf raft	floor go	going floor	stealing	Persia...
+	-	رو	<E>	انداختن	<E>	-	-	رو انداختن	ro andakht	fac e throw	throwing face	asking	Persia...
+	-	رو	<E>	داشتن	<E>	-	-	رو داشتن	ro dasht	face have	having face	being ...	Persia...
+	-	سر	<E>	آوردن	<E>	-	+	سر در آوردن	sar daravard	head evolve	evolve head	Unders...	Persia...
+	-	سماق	<E>	مکیدن	<E>	-	+	سماق مکیدن	somagh mekid	sumac suck	sucking sumac	wating...	Persia...
+	-	سنگ	<E>	انداختن	<E>	+	+	سنگ انداختن	sang andakht	stone throw	throwing s...	causin...	Persia...
+	-	سوسه	<E>	آمدن	<E>	-	-	سوسه آمدن	sose amad	weevil come	coming weevil	causin...	Persia...
+	-	صغرا کبرا	<E>	چیدن	<E>	-	-	صغرا کبرا چیدن	soghra kobra chid	Cobra Soghra...	Picking Co...	not ta...	Persia...
+	-	غاز	<E>	چراندن	<E>	-	+	غاز چراندن	ghaz charand	Goose graz	grazing goos	being ...	Persia...
+	-	قاپ	<E>	زدن	<E>	-	-	قاپ زدن	ghap zad	rogue hit	hitting rogue	sudden...	Persia...
+	-	قار قار	<E>	کردن	<E>	-	+	قار قار کردن	ghar ghar kard	caw caw do	doing caw caw	being ...	Persia...
+	-	قپی	<E>	آمدن	<E>	-	-	قپی آمدن	ghopi amad	boast come	coming boast	Jactit...	Persia...
+	-	قمپز	<E>	کردن	<E>	-	-	قمپز در کردن	gompoz darkard	bluff shoot	shooting b...	a pers...	Persia...
+	-	لب	<E>	چیدن	<E>	-	-	لب برچیدن	lab barchid	lip remove	removing lip	being ...	Persia...
+	-	لب	<E>	گزیدن	<E>	-	+	لب گزیدن	lab gazid	lip bit	biting lip	feel s...	Persia...
+	-	لگا لگا	<E>	کردن	<E>	-	-	لگا لگا کردن	lek lek kard	demoiselle do	doing demo...	doing ...	Persia...
+	-	لنگ	<E>	انداختن	<E>	-	+	لنگ انداختن	long andakht	textile throw	throwing t...	give up	Persia...
+	-	لیچار	<E>	گفتن	<E>	+	-	لیچار گفتن	lychar goft	lychar say	lychar saying	trash ...	Persia...
+	-	م	<E>	دادن	<E>	-	+	م با دادن	nam pas nadad	Humidity giv...	giving bac...	being ...	Persia...
-	+	آب	<E>	کردن	<E>	-	+	آب باز کردن	ab baz kard	water open	opening water	happen...	Persia...
-	+	آب	<E>	برداشتن	<E>	-	+	آب برداشتن	ab bardasht	water pickin...	pick up water	Being ...	Persia...
+	-	آب	<E>	دادن	<E>	-	+	آب دادن	ab dad	water give	giving water	reachi...	Persia...
-	+	آب	<E>	دیدن	<E>	-	+	آب دیدن	ab did	water see	seeing water	Damage	Persia...

Figure 2. Unicode database

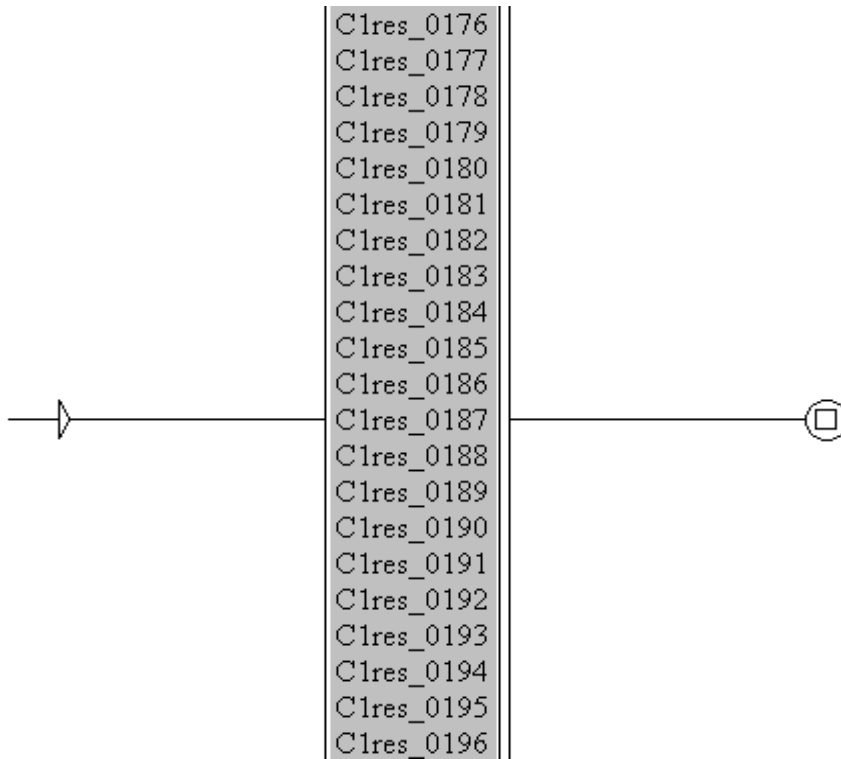


Figure 3. C1res.grf

The graph of Figure 4 illustrates the expression of example (2a) and (2b) that does not allow for the passive so we can see that in the output only the transitions corresponding to the passive will be produced.

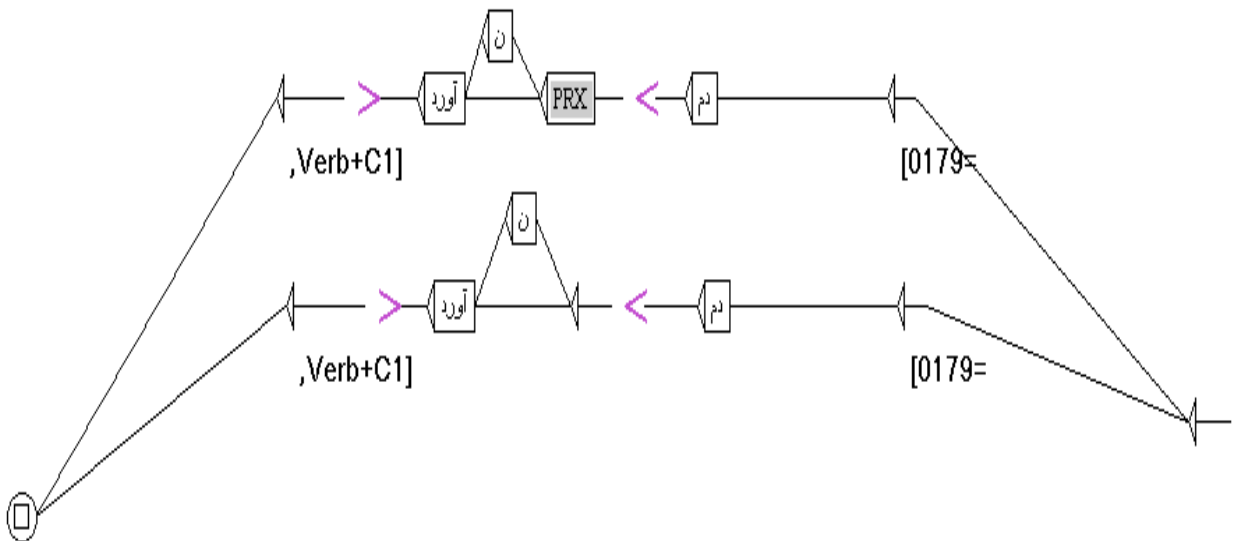


Figure 4. C1res.0179.grf

On the other hand, in Figure 5, we have expression example of (1a) and (1b) that allows for the passive, and the corresponding paths are in fact produced in the output.

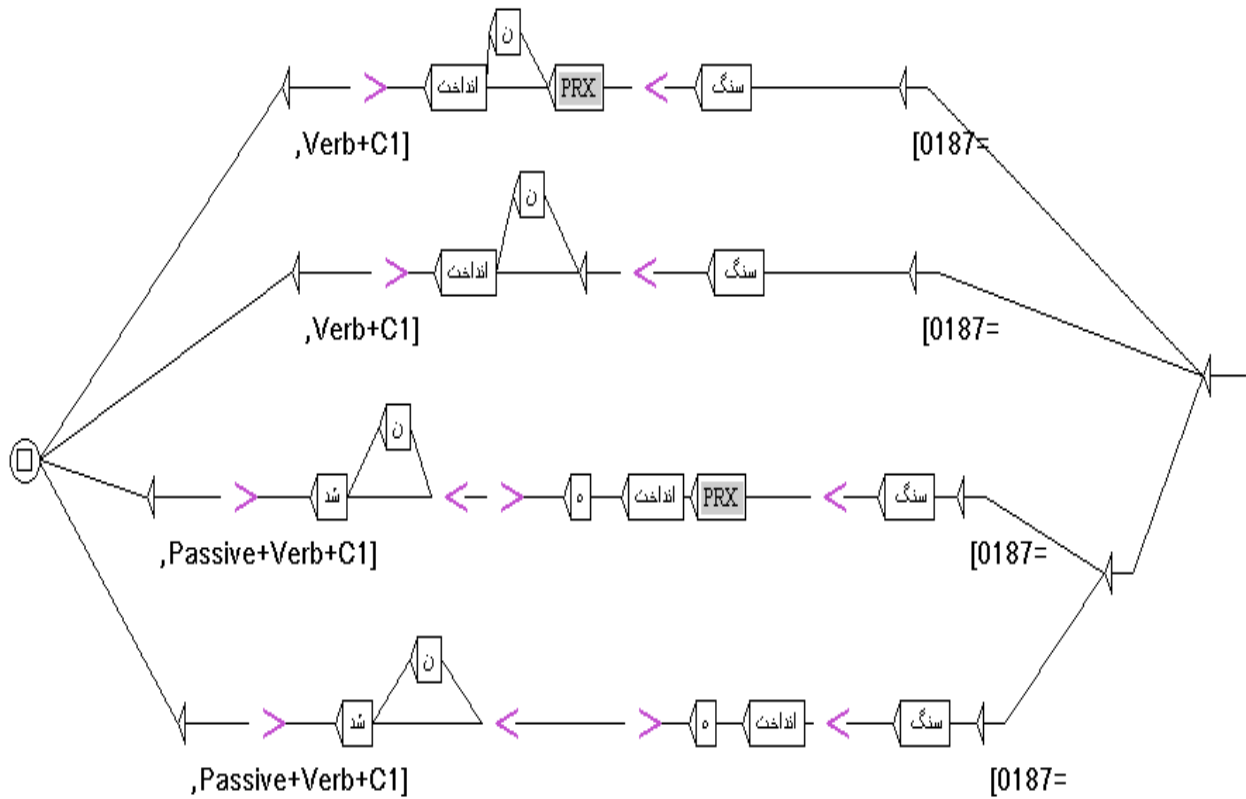


Figure 5. C1res.0187.grf

We can now use the resulting graph to look for idioms in the corpus. Figure 6 is an excerpt of the concordance from the idioms found in the TEP corpus. At this stage, only 268 matches were found. In the concordance some sentences, which captured, do not have the idiomatic meaning, which means that, they have the literal meaning of that particular paragraph that sounds normal.

م . 100701 اما مایل نیستم که با سرو گوش [\[Verb+C1, دادن=0002\]](#) زیادی برای همه مسئله درست کنم .  
 این به شما علامت میدم . 462879 اینا با [\[Verb+C1, گرفتن=0005\]](#) دود غلیظ سیاهی از خودتون میدن  
 شما چی فکر میکنید که این اتفاق به وسیله [\[Verb+C1, گرفتن=0005\]](#) یک صفحه رخ داده بشه . 353933 م  
 219165 من دوستام را دیدم که جلوی چشم [\[Verb+C1, گرفتن=0005\]](#) 219166 کجا تو فایق . 219167  
 د علفهای زیر شروونی که تو مزرعه بود ن [\[Verb+C1, گرفتن=0005\]](#) 127736 اون هم سال 1935 بود .  
 نه عزیزم تو هنوز هم شایسته ای . 322361 [\[Verb+C1, کردن=0013\]](#) فقط یک امتیاز اضافیه . 322362 ت  
 بس کن هری سازمان کسایی را که عمه شون را [\[Verb+C1, کردن=0013\]](#) نفیفرسته آزکابان . 24428 از طرف  
 . 63289 گمش نکنی . 63290 من منظورشو از [\[Verb+C1, کردن=0013\]](#) مارچ نفهمیدم من فقط . 63291 کنت  
 ار دوست داری ای . 99404 اینویکه پسر خوب [\[Verb+C1, آوردن=0014\]](#) تا همیشه فروتنانه سفارش بده خی  
 ریدم . 205391 ولی بین چه افتضاحی به [\[Verb+C1, آوردن=0014\]](#) 205392 چی از واگن ها دزد  
 8 نه ببخشید مامان من یک عقیده محکم برای [\[Verb+C1, ندادن=0024\]](#) دارم . 148609 اگر نیتونی بری ت  
 نگ هی حمله کنیم . 543588 جانگ بوگ آماده [\[Verb+C1, آوردن=0025\]](#) خسارات بزرگه . 543589 این شما  
 درت شنوایی و نه قدرت روانی . 87287 برای [\[Verb+C1, آوردن=0025\]](#) برابر قدرت با عظمت صدای واقعی  
 46407 کمک رمی کمک . 46408 امیل شروع به [\[Verb+C1, خوردن=0026\]](#) با لامپ . 46409 امیل به طرف من  
 اهزهایی که خودتون را به جایبرنس موکانگ [\[Verb+C1, زدن=0031\]](#) 424958 همه جا هستند . 424959 بل  
 اغش به پدر من . 233216 میدونستم که برای [\[Verb+C1, زدن=0031\]](#) قطعات باهوشیم . 233217 این کابریو  
 اریس . 16022 دنبال من بیان لطفا . 16023 [\[Verb+C1, گرفتن=0032\]](#) این دو تا هرم . 16024 خیلی نادر  
 0 همش کار اوناس این سوسیالیتها . 227701 [\[Verb+C1, زدن=0049\]](#) به همه چی . 227702 میخواستم بهت  
 8 همش کار اوناس این سوسیالیتها . 271269 [\[Verb+C1, زدن=0049\]](#) به همه چی . 271270 می خواستم بهت  
 جور ادمی هستی . 339664 چطوری با یکه جور [\[Verb+C1, گند=0049\]](#) 339665 گند کاریت را قبول کن و  
 م رفیق من فقط نموش کردم . 289371 من از [\[Verb+C1, زدن=0049\]](#) دائم خسته و بیمارم . 289372 اون  
 رداون چقدر ساکنه . 345957 خیلی از مردم [\[Verb+C1, زدن=0049\]](#) بهش عالیه . 345958 آقای وزیر . 3

Figure 6. Concordance

### 3.4. Corpus parsing and idioms candidate extraction

The corpus<sup>6</sup> that we will use for this thesis project is a Persian movies subtitle which, we get it from the natural language processing department of Tehran University and it has 556,234 sentence delimiters, 15,166,987 (64,492 diff) tokens, 4,485,147 (64,365) simple forms, 3,239,250 (10) digits. In this project, all idioms that have a *NO CI V* structure will be captured from the chosen websites and dictionaries. Then they will be applied in the Persian folder of Unitex. The formal description will take the form of a database where the lexical distributional and transformational properties of idioms will be encoded.

<sup>6</sup> <http://ece.ut.ac.ir/NLP/resources.htm>

## 4. Evaluation

As the purpose of this dissertation project is to build a lexical database, the evaluation will consist in three tasks, which are; (i) estimating the scope of the database; (ii) determining the precision of the task of identifying the idioms, using the Unitex tools; (iii) association measures.

### 4.1. Estimation of idioms database scope

To estimate the scope of the database idioms, we produced a dummy FST (Figure 7) that would extract all the instances containing the verb<sup>7</sup> and the lexical constant C1 from the text, allowing for a window of length 0 to 5 words (that is observable in the gray box below), which inserted between the verb and the constant. In this dummy FST, the variables introduced by the symbol “@” correspond to the columns of the Table 4, where @F= past tense tensehe verbs, @G= present tense of the verb, @I= property of accepting passive voice or not, an exactly other variables are exactly similar to Figure 1. By intersecting the database with this FST, we were able to produce a set of transducers that extract all instances that “*potentially*” include the idioms of the database (Figure 8).

Table 4. Database with the past and present stem of each idiom verb

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
NO=:Hum	NO=:Nhum	C1	Det	Verb	past tense	present ten	Modify	Passive	Literal	Persian idioms	Translitera tion	Gloss	Rough translation	Exact meaning	Reference
-	+	برق	<E>	افتادن	افتاد	افت	<E>	-	-	برق افتادن	bargh oftad	electricity fall	falling electricity	being too shiny and clean	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	افلاس	<E>	افتادن	افتاد	افت	<E>	-	-	افلاس افتادن	eflas oftad	Bankruptcy fall	falling bankruptcy	Getting caught in the poverty	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-	پس	<E>	افتادن	افتاد	افت	<E>	-	-	پس افتادن	pas oftad	back fall	falling back	fainting	Persian proverbs and idiomatic expressions book
+	-	تک	<E>	افتادن	افتاد	افت	<E>	-	-	تک افتادن	tak oftad	monad drop	dropping monad	being alone	Persian slang dictionary

<sup>7</sup> Since we do not have the dictionary of the inflected forms for all the verbs in the database, we manually added the stem of past and present (e.g. کرد, kard, کن, kon “to do” and زد, zad, زن, zan “to hit”) which is accessible in appendix.

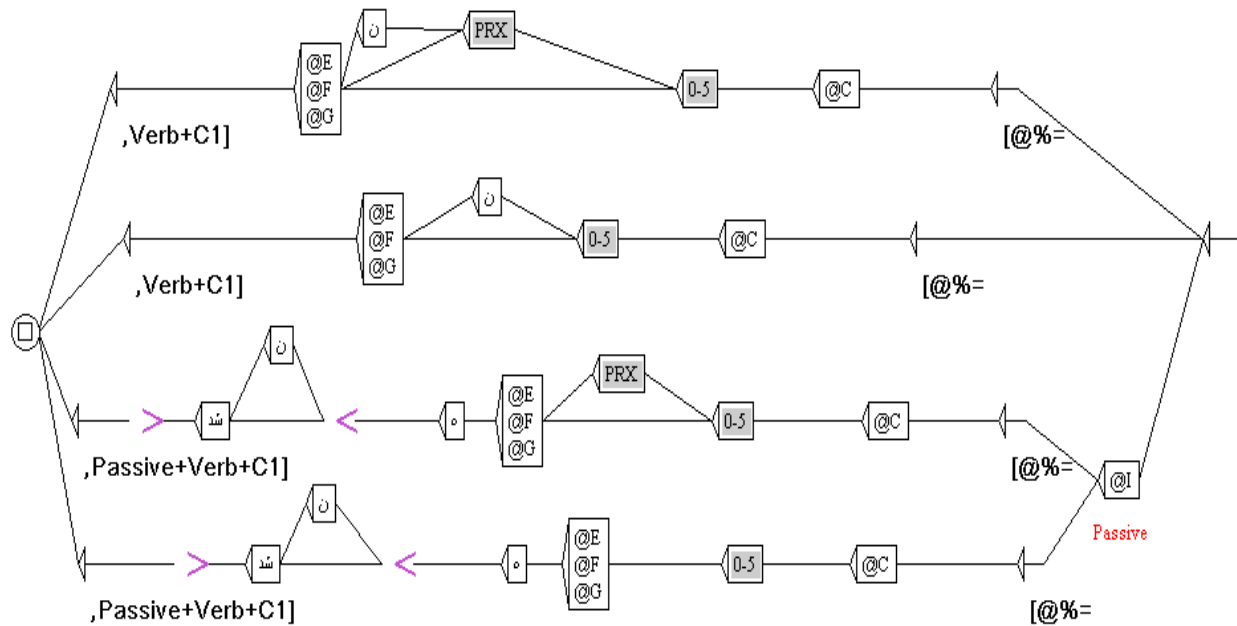


Figure 7. Extract.ref.grf

م . 100701 اما مایل نیستم که با سرو گوش [Verb+C1, دادن=0002] زیادی برای همه مسئله درست کنم .  
 انواده ای که حکومت ازش می‌گه . 7801 اونا [Verb+C1, تاش دادن=0002] اونا شیطونی نشونش میدان .  
 تیش درست کنیم . 95581 می‌خواستیم با بصرم [Verb+C1, آتش درست کردن=0004] را تیرین کنیم . 95582 باشه  
 هیم کرد . 311825 آماده شو برای انفجار و [Verb+C1, آتش روشن کردن=0004] هون جوری که گفتیم . 311826  
 این به شما علامت میدم . 462879 اینا با [Verb+C1, آتش گرفتن=0005] دود غلیظ سیاهی از خودشون میدان  
 شما چی فکر میکنید که این اتفاق به وسیله [Verb+C1, آتش گرفتن=0005] بگه صفحه رخ داده بشه . 353933 م  
 219165 من دوستانم را دیدم که جلوی چشمم [Verb+C1, آتش گرفتن=0005] . 219166 کجا توفایق . 219167  
 د علفهای زیر شیره‌ونی که تو مزرعه بودی [Verb+C1, آتش گرفتن=0005] . 127736 اون هم سال 1935 بود .  
 مغزت باید بیدار بشه . 87911 اونا من و [Verb+C1, آتش گرفتن=0005] . 87912 کدوم کیر  
 نه عزیزم تو هنوز هم شایسته ای . 322361 [Verb+C1, آتش گرفتن=0005] فقط یک امتیاز اضافیه . 322362  
 بس کن هری سازمان کسایی را که عمه شون را [Verb+C1, آتش گرفتن=0005] نفیفرسته آژکابان . 24428 از طرف  
 . 63289 گمش نکنی . 63290 من منظورشو از [Verb+C1, آتش گرفتن=0005] مارچ نفهمیدم من فقط . 63291 کنت  
 ر متوجه جادویی فوراً اعزام شدند . 63215 [Verb+C1, آتش گرفتن=0005] و حافظه اش را اصلاح  
 نفرم که پیام . 82599 خوبین بر شما سلامت [Verb+C1, آتش گرفتن=0005] بهتر است . 82600 امروز میخو  
 از دوست داری ای . 99404 اینو بگه بصر خوب [Verb+C1, آتش گرفتن=0005] تا همیشه فروتنانه سفارش بده خی  
 ریدم . 205391 ولی بین چه افتضاحی به [Verb+C1, آتش گرفتن=0005] . 205392 چی از واگن ها دزد  
 دتاون با ورود به این محاکمه . 8950 چند [Verb+C1, آتش گرفتن=0005] . 8951 د  
 نگه می‌کنیم . 543588 جانگ بوگ آماده [Verb+C1, آتش گرفتن=0005] خسارات بزرگه . 543589 این شما  
 درت شنوایی و نه قدرت روانی . 87287 برای [Verb+C1, آتش گرفتن=0005] برابر قدرت با عظمت مدای واقعی  
 46407 کمک رمی کمک . 46408 امیل شروع به [Verb+C1, آتش گرفتن=0005] با لامپ . 46409 امیل به طرف من  
 م تأثیر میگذارند . 363938 تو مزه ترشی و [Verb+C1, آتش گرفتن=0005] شیرینی خنثی  
 ز بنگشت . 467912 من مطمئنم اونا شد به [Verb+C1, آتش گرفتن=0005] . 467913 اگه نه این جزیه کس.

Figure 8. Concordance of dummy FST

Figure 8 illustrates a concordance of the matched strings. In this concordance, one can see the idioms delimited by [], the number of idioms and its type (Verb +C1).

As we want to have a more precise perspective of what is going on in the data, we selected two verbs; one is the most frequent verb کردن *kardan* (to do), which can also function as a support verb, and another a full verb زدن *zadan* (to hit). We prepared a small dictionary of these verbs' inflected forms in order to capture all idioms involving these verbs. After that, we selected two sets of idioms (those that had کردن *kardan* (to do), and زدن *zadan* (to hit) as their verbs) that corresponds to about 30% of entire database entries. In the database, there are 57 idioms with کردن *kardan* (to do) that is, 15.6% of the entire database entries, and 47 with زدن *zadan* (to hit), 12.9% of the whole database. Afterwards, we built two FSTs, for each Verb-C1 combination; one with a window of 0 to 5 inserted words, and another one without any insertion between verb and lexical constant C1 (Figures 9 to 12).

For the verb کردن *kardan* (to do) FST with the insertion (Figure 9), we obtained 510 matches, which correspond to 36 different idioms (types).

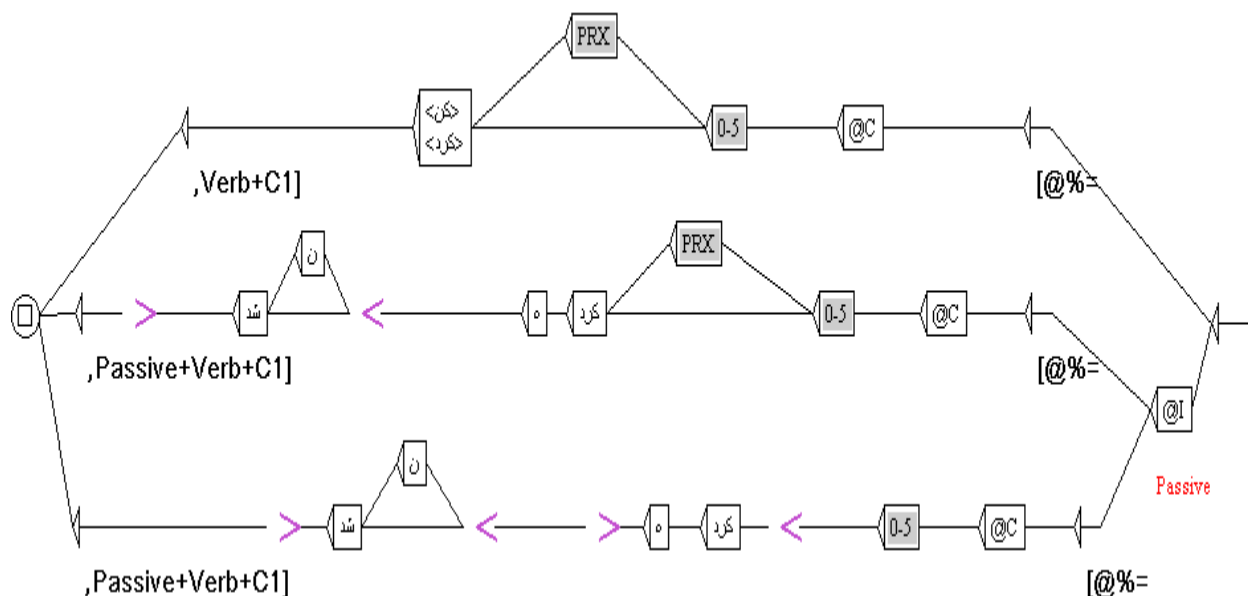


Figure 9. Dummy reference FST for idioms with the verb “کردن *kardan* (to do)” with 0-5 insertion window

For the verb کردن *kardan* (to do) FST without insertion (Figure 10), we obtained 107 matches, which correspond to 23 different idioms (types).

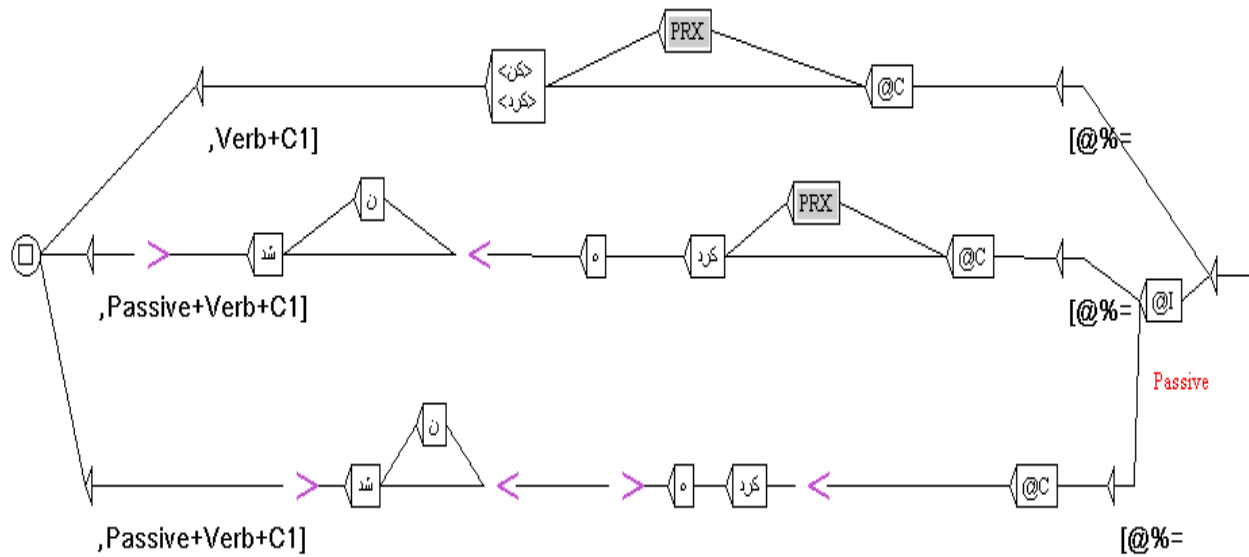


Figure 10. Reference FST for idioms with the verb “کردن *kardan* (to do)” without any insertion

For the verb زدن *zadan* (to hit), FST with the insertion (Figure 11), we obtained 84 matches, which correspond to 12 different idioms (types).

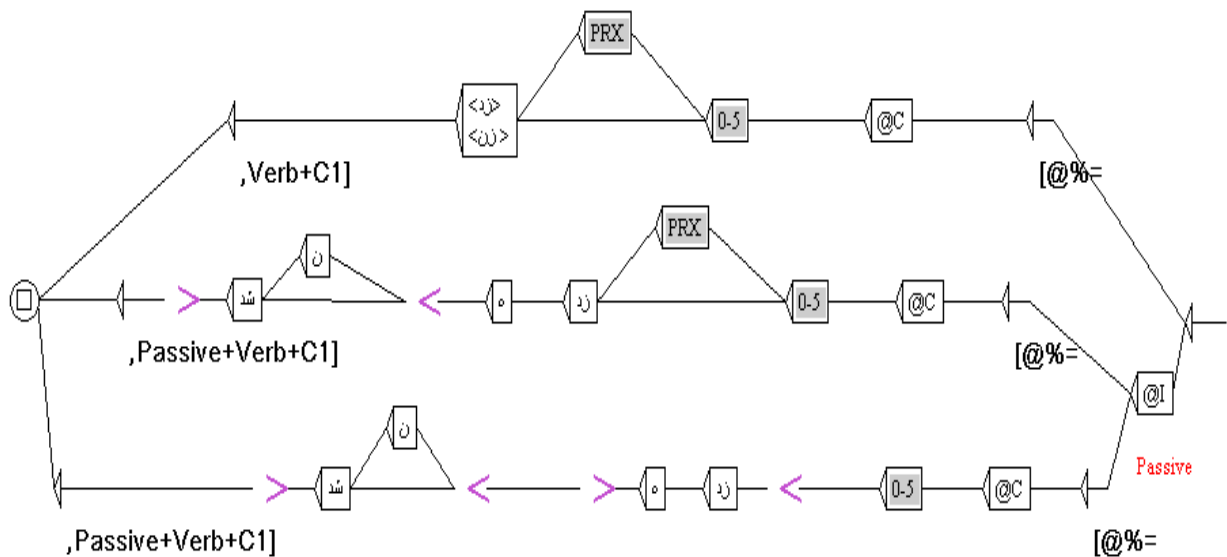


Figure 11. Dummy reference FST for idioms with the verb “زدن *zadan* (to hit),” with 0-5 insertion window

Finally, for the verb *زدن* *zadan* (to hit), FST without insertions (Figure 12), we obtained 60 matches, which correspond to seven different idioms (types).

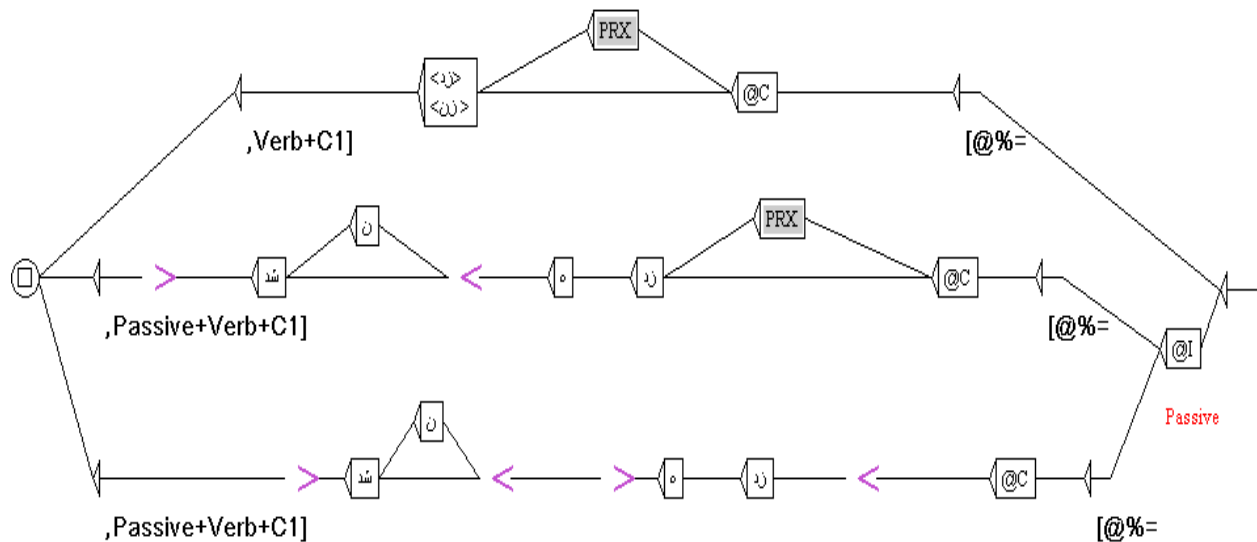


Figure 12. Reference FST for idioms with the verb “*زدن*, *zadan* (to hit),” without any insertion

The breakdown of these results is shown in the following table (Table 5).

Table 5. Experiments with the idioms involving verbs *کردن* *kardan* (to do) and *زدن* *zadan* (to hit)

Experiment	Number of entries (types)	Number of matches	Different idioms (types)
Entire database with insertion	364	1,754	115
Entire database without insertion	364	584	68
<i>کردن</i> <i>kardan</i> (to do) with insertion	57	510	36
<i>کردن</i> <i>kardan</i> (to do) without insertion	57	107	23
<i>زدن</i> <i>zadan</i> (to hit) with insertion	47	84	12
<i>زدن</i> <i>zadan</i> (to hit) without insertion	47	60	7

The first overall conclusion to be drawn from Table 5 is that, apparently, the insertion window increased the total number of matches: 3 times more than without the insertion in the case of the *entire database*; 4.8 times more than without the insertion in the case of the idioms with *کردن* *kardan* (to do) and 1.4 times more than without the insertion for the idioms with *زدن* *zadan* (to

*hit*) The difference is much more relevant for the two sets of idioms with the verbs کردن *kardan* (to do) and زدن *zadan* (to hit). In the case of کردن *kardan* (to do), there is a drop 80% in the number of matches, but only about 29% in the case of زدن *zadan* (to hit) (though the number of entries is small). This means that, depending on the verb, the scope of the window may have to be empirically determined and adjusted accordingly, in order to capture the idioms better. Moreover, the number of different idioms captured by the graph without these insertions is significantly less: 31.6 % of the idioms in the database are captured by the graphs with the insertions while only 18.7% are matched when the insertions are removed; and in the case of کردن *kardan* (to do) and زدن *zadan* (to hit), the percentage of different idioms that are captured by the graph with these insertions are respectively 63% and 25%, while without the insertions are only 40% and 15%.

Then, we went manually through the resulting concordances of each dummy FSTs, with and without insertions, and assigned each match with one of these three following tags:

- 1) The idiom is present and has been matched by the FST: *true-positive* (TP)
- 2) The elements of the idiom are present, but they do not have an idiomatic value: *false-positive* (FP)
- 3) The elements of the idioms are present, but there is no relation between them: *spurious co-occurrence* (S)

The results are given in the following table (Table 6).

Table 6. Concordances result of each dummy FSTs

	کردن <i>kardan</i> (to do)			زدن <i>zadan</i> (to hit)		
	TP	FP	S	TP	FP	S
FST w/ Ins	34	81	395	49	18	17
FST w/o Ins	53	54	0	42	18	0

It is noteworthy that the number of spurious matches in the verb کردن *kardan* (to do) with insertions is almost 5 times more than the number of false positive matches, while with the verb زدن *zadan* (to hit) the number of spurious matches is much smaller. This difference is likely to

be related to auxiliary status of کردن *kardan* (to do) in Persian (as a support verb, for example), while زدن *zadan* (to hit), is mostly used as a full verb. This, once again, indicates that each verb may require a different strategy in view of achieving better recall.

In the case of a recall, the FST for کردن *kardan* (to do) with the insertions, by chance captured 3 idioms that not included in the database but, involve some nouns which, present in the list of idioms (examples 1 to 3);

- (1) Persian: کف دستمو بو نکرده  
 Transliteration: kafe dastamo bo *nakarde*  
 Morph<sup>8</sup>: palm of hand my smell did not he/she  
 Translation: he/she did not smell my palm  
 Gloss: he/she did not know anything about something
- (2) Persian: آتش درونم رو خاموش کنم  
 Transliteration: atashe daronam ro<sup>9</sup> khamosh *konam*  
 Morph: fire inside my turn off to do I  
 Translation: I turn off my inside fire  
 Gloss: I am trying to make myself calm and relax
- (3) Persian: تخم پیدا کنی  
 Transliteration: tokhm peyda *koni*  
 Morph: egg finds to do  
 Translation: finding egg  
 Gloss: becoming brave

In the first example the verb “*kardan*”, happens to co-occur with the noun کف *kaf* (floor), with which it also forms the idiom کف کرد *kaf kard* (to be shocked). However کف دستم *kafe dastam* (palm hand -my) is a compound word (hand palm), *kafe dast*, formed by کف *kaf* (floor) and دست *dast* (hand). This idiom, thus, corresponds to another formal class involving two compliments or

<sup>8</sup> Morphological reading of each word.

<sup>9</sup> In Persian, a direct object is marked with the post-positioned particle *rā* را or *ro* رو which immediately follows the noun. for example: از کتابفروشی کتاب را خریدم, az ketābforuši ketāb rā kharidam, from bookstore book \_ra bought-I “I bought the book from (a/the/some) book store.”

rather a support verb, predicative noun and a dative compliment, which just happens to overlap the simple idiom. A similar situation occurs with the other example (number 2). In the third example, we find the support verb construction *پیدا کردن* *peyda kardan* (to do finding) with the object *تخم* *tokhm* (egg), which means, “to become brave” while, the idiom in the table *تخم کردن* *tokhm kardan* (to do an egg) means, “to be very frightened”. These instances let one to think that many more idioms may still be uncharted and call for the extension of this work to other idiom classes, with different syntactic structures. It calls for the problem of disambiguating between idioms involving similar lexical material

In short, when one considers a full verb like *زدن* *zadan* (to hit), since all idioms captured by the FSTs were in the database, opening the insertion window did not diminish the precision significantly. On the other hand, for an auxiliary verb such as *کردن* *kardan* (to do) this leads to a large drop in precision. Naturally, these findings cannot be directly extrapolated to the remainder of the database, and a systematic survey of all the verbs involved in these V+C1 structures should be verified. That means; the study should still be made not only for *کردن* *kardan* (to do) and *زدن* *zadan* (to hit), but also for the remainder of the verbs.

#### 4.2. Determine precision of identification by using the Unitex tools

By precision, we mean to try and measure how many of the idiomatic expressions retrieved from the corpus were correctly captured. In other words, we defined precision as  $TP / (TP + FP)$ , ignoring spurious matches (S). In view of these data, we calculated the precision of each group of idioms (Table 7).

Table 7. Precision

	<i>کردن</i> <i>kardan</i> (to do)	<i>زدن</i> <i>zadan</i> (to hit)
FST w/ Ins	0.29	0.73
FST w/o Ins	0.50	0.70

Table 7 indicates that, for idiomatic combinations with the full verb *زدن* *zadan* (to hit), the finite state methods that were used here can produce relatively good results; while for more grammaticalized verbs, such as the verb *کردن* *kardan* (to do), which often functions as a support verb, the form of variation, namely the length of the insertions, can dramatically reduce the

precision. This is linked to the fact that many of these idioms are ambiguous between a literal and an idiomatic reading. Besides, the small reduction in the precision of the FST without the insertion with verb *زین zadan (to hit)* results from a lower number of idioms that are being retrieved when compared to *کردن kardan (to do)*.

### 4.3. Association measures

For the idioms that were captured by the FST using the entire lexicon-grammar, we calculated two association measures; t-test<sup>10</sup> and chi-square<sup>11</sup>. For the estimation of t-test and chi-square as applied to idioms identification, we adopted the proposal of Manning and Schütze (2005), chapter 5. Table 8 and Table 9 shows the values of t-test and chi-square for the idioms with *کردن kardan (to do)* and *زین zadan (to hit)*. On the two rightmost columns, we provide a binary value for the comparison of the t-test and chi-square values against the reference statistical relevance threshold (P).

Table 8. "کردن, kardan (to do)" T-test and chi-square

Idiom	Example (transliteration) translation/gloss	T-test	Chi-square	t > p	$\chi^2 > p$
[0001]	آتش کردن (atash kardan) shooting	-9.36316	8.575242	0	1
[0003]	باد در کردن (bad darkardan) farting	2.629322	21.82092	1	1
[0005]	تابلو کردن (tablo kardan) revealing something	0.641734	1.164957	0	0
[0007]	تروش کردن (toesh kardan) becoming angry	1.78106	29.36749	0	1
[0008]	جفت کردن (joft kardan) fearing	-0.96383	0.479417	0	0
[0010]	کف کردن (kaf kardan) getting excited	-0.43306	0.132633	0	0
[0012]	باد کردن (bad kardan) getting pregnant	2.629322	21.82092	1	1
[0015]	قاطی کردن (ghati kardan) becoming angry	2.951347	544.3165	1	1
[0016]	کلید کردن (kilid kardan) being bête noire	0.734618	2.060893	0	0
[0017]	تخم کردن (tokhm kardan) greatly fear	-1.5742	0.975672	0	0
[0018]	ردیف کردن (radif kardan) making everything alright	1.311004	23.86749	0	1
[0021]	پف کردن (pof kardan) feeling self-respect	2.357399	149.8065	0	1
[0022]	تروش کردن (torsh kardan) huffing	1.78106	29.36749	0	1
[0027]	پنچر کردن (panchar kardan) being too tired	1.292239	19.62157	0	1
[0036]	آب باز کردن (ab bazkardan) happening something wrong	-5.20311	9.252444	0	1
[0038]	افاقه کردن (efaghe kardan) having good effect	0.986731	74.36311	0	1
[0040]	بچه درست کردن (bache dorostkardan) reproduction	-52.302	52.0676	0	1
[0041]	درز کردن (darz kardan) make known something private	0.814233	3.616851	0	0
[0044]	غو غا کردن (ghogha kardan) doing the best	0.907116	8.978188	0	1
[0046]	کولاک کردن (kolak kardan) doing the best	1.357918	46.94466	0	1
[0049]	داغ کردن (dagh kardan) becoming angry	0.938472	1.681747	0	0
[0055]	کف کردن (kaf kardan) being tired	-0.43306	0.132633	0	0
[0056]	قفل کردن (ghofl kardan) being shocked	4.544733	192.5728	1	1

<sup>10</sup> [http://en.wikipedia.org/wiki/Student's\\_t-test](http://en.wikipedia.org/wiki/Student's_t-test)

<sup>11</sup> [http://en.wikipedia.org/wiki/Chi-squared\\_test](http://en.wikipedia.org/wiki/Chi-squared_test)

Table 9." زدن *zadan* (to hit)."T-test and chi-square

Idiom	Example (transliteration) translation/gloss	T-test	Chi-square	t > p	$\chi^2 > p$
[0002]	جاذین (ja <i>zadan</i> ) fearing to do something	2.949175	30.63058	1	1
[0004]	گند زین (gand <i>zadan</i> ) sabotage	5.590053	2651.799	1	1
[0016]	قازین (ghat <i>zadan</i> ) becoming angry	1.402532	238.6465	0	1
[0023]	پرسه زین (parse <i>zadan</i> ) walking without a goal	0.896753	7.804938	0	1
[0026]	نیش زین (nish <i>zadan</i> ) Squibbing	1.970058	259.7837	0	1
[0035]	قاپ زین (ghap <i>zadan</i> ) sudden stealing	1.405453	319.5216	0	1
[0045]	جوش زین (josh <i>zadan</i> ) be concerned	1.288642	18.74106	0	1

Table 10. Comparison of t-test and chi-square

	$T \& \chi^2$ (1 1)	$\sim T \& \chi^2$ (0 1)	$T \& \sim \chi^2$ (1 0)	$\sim(T \& \chi^2)$ (0 0)
کردن ( <i>kardan</i> ) to do	4	11	–	8
زدن ( <i>zadan</i> ) to hit	2	5	–	–

These results are summarized in Table 10. As one can see, there are 8 idiom types with کردن *kardan* (to do) that were not captured by any of this association measures, which corresponds to 35% of the entries with this verb (this also corresponds to 27% percent of both verbs). Secondly, t-test only captures about 27% of the idioms (and always when chi-square also does it), while chi-square is able to capture 73% of all the idioms. The results of these two association measures indicate that chi-square is a better performing measure than t-test for these types of idioms. Still, for the verb کردن *kardan* (to do) only 59% were captured which is above the precision one gets with the FST without the insertion, while for زدن *zadan* (to hit) all idioms types are captured. Thus, we can conclude that the association measures are valid strategies to capture this kind of idioms; however, depending on the verb that is at stake, namely the fact that one of the verbs is a support verbs, the results may vary.

## 5. Conclusion

The initial objectives of this dissertation were achieved; we have made a selection of web sources by using Google and other browsers that are powered by Google, such as AVG, and we have built a database that contains 364 (Verb+C1) idiomatic expressions. We have built a set of FSTs in order to explore Persian texts, and tested them using a corpus. We have retrieved a Persian corpus consisting of movies subtitles, which contains 556,234 sentences, 15,166,987 (64,492 diff.) tokens, and 4,485,147 (64,365) simple forms. We have been able to extract idioms candidates by using these FSTs and we produced an evaluation of these results, both in terms of precision of these FSTs, and by comparing them with the use association measures (t-test and chi-square). Results indicate that depending on the verb, namely, if the verb has a more grammatical status or is a full verb, it may be necessary to adjust insertion windows between the key elements of the idiom; precision also varies depending on the verb. Association measures are valid strategies to capture these types of idioms; however, their precision depends on the idioms' verb, as the presence of support verbs may vary the results.

Of course, much work still left to be done; for instance:

- Enhance the syntactic complexity of the reference graphs in order to capture other syntactic patterns such as permutations that typically occur in the spoken or informal speech (Persian verbs have special inflection paradigms for informal, spoken forms; Bahareh Kakanaeeni, *in progress*); e.g. وحید نیش زد (N0 C1 V), *Vahid nish zad*, Vahid sting hit, “Vahid hit the sting”, Vahid is taunting. In the spoken form, this idiom allows for the permutation of the verb and the frozen complement, yielding N0 V C1; e.g. وحید زد نیش رو, *Vahid zad nish ro*, Vahid hit sting, Vahid hit the sting, Vahid is taunting.
- Plug the existing graphs with the Persian dictionary of inflected forms, which is currently being developed (Bahareh Kakanaeeni, *in progress*); this could improve the precision in general, and enable us to use association measures to capture idiom candidates directly from the corpus;

- Extend the collection of idioms by using other sources, augmenting the lexical coverage of the lexicon-grammar of the idioms.
- Extend the collection of idioms to other types of Persian idioms; from the survey we did on our lexical sources, we have already found many other idioms which can easily be integrated in Maurice Groos (1996, page 249), formal classification of frozen sentences, namely;

CP1    N0 Prep C1 V

CNP2    N0 Prep C2 N1 V

These two structures are illustrated below;

وحدید با دنده سنگین رفت

Vahid ba dandei sangin raft,

Vahid with low gear went,

Vahid went with low gear,

Vahid went slowly

وحدید زیر اب امی را زد

Vahid zirabe Emmy ra zad,

Vahid under\_water Emmy hit,

Vahid hit Emmys underwater.

Vahid cheated on Emmy.

## References

### Active bibliography

- Aisenstadt, Elbert. 1979. *Collocability restrictions in dictionaries*. In R. R. K. Hartmann, D. Dictionaries and their users, 71-74. Exeter Linguistic Studies.
- Baker, M. (1992) *In Other Words*. London: Routledge.
- Benson, M., Benson, E. & Ilson, R. (1990). *Collocations and general-purpose dictionaries*. *International Journal of Lexicography*, 3 (1), 23-35.
- Benson, M., Benson, E. & Ilson, R. 1986a. *Lexicographic Description of English*. Amsterdam and Philadelphia: John Benjamin.
- Cain, K. & Oakhill, J. & Lemmon, K. (2004). *The Relation between Children's Reading Comprehension Level and Their Comprehension of Idioms*. *Journal of Experimental Child Psychology*, 90, 65-87.
- Carter, R (1998). *Vocabulary: applied linguistics perspectives*, second edition, London/ New York: Routledge
- Chang & Fischer. (2006). *Understanding idioms*. Retrieved January 22, 2010 from <http://www.icsi.berkeley.edu/~nchang/pubs>.
- Collins, Cobuild. *Dictionary of Idioms*. (2000) By Collins <http://www.elearnaid.com/>
- Cruse, David Alan. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.
- Crystal, David. (1985). *A dictionary of linguistics and phonetics*. Second edition. Oxford: Basil Blackwell.
- Curry, D. (1994) *Illustration American idioms*, published by The Materials Branch English Language Programs Division, United States Information Agency, Washington DC.
- Fazly, A. & S. Stevenson. (2006). *Automatically constructing a lexicon of verb phrase idiomatic combinations*. Retrieved January 19, 2010 from:<http://acl ldc.upenn.edu/E/E06-1043>. Pdf.
- Fernando, C. (1994). *Idioms and Idiomaticity*. Oxford: Oxford University Press.
- Glaser, R. (1984): *Terminological Problems in Linguistics with Special Reference to Neologisms*, in: Exeter 83 Proceedings (Lexicographical Series Major 1) Ed. By R.R.K. Hartmann, Tübingen: Maxx Niemeyer, 345-351.
- Hatch, E and Brown, C. (1995). *Vocabulary, Semantics, and Language Education*, Cambridge University Press.

- Hill, J. (2000) *Revising Priorities: From Grammatical failure to collocational success*. In Lewis M. (Ed). *Teaching Collocations* (pg. 47 - 67). Thompson Hienle. Hove: UK
- Howarth P. (1998) *the Phraseology of Learners' Academic Writing*. In Cowie A.P. (ed.) *Phraseology: Theory, Analysis, and Applications*. Oxford: OUP, 161-186
- Larson, M.L. (1984). *Meaning Based Translation: A Guide to Cross Language Equivalence*. London and New York: University Press of America
- Lewis, M. 2000. *Teaching collocations*. London: Language Teaching Publications.
- Manning C. & Schütze H (2005), *Foundations of Statistical Natural Language Processing*, MIT Press, Cambridge, Massachusetts.
- Maurice Groos. (1996). *Lexicon-Grammar*. In Brown, K. & Miller, J. (eds.), *Concise Encyclopedia of Syntactic Theories*. Cambridge: Pergamon, pp. 244–259.
- McCarthy, M. (1998): *Spoken Language and Applied Linguistics*, Cambridge: CUP.
- McMordie, J. S. (1983). *English Idioms and How to Use Them*. Moscow: Vyschaja Shkola.
- Moon, R. (1997): *Vocabulary Connections: Multi-Words Items in English*.
- Moon, R. (1998) *Fixed Expressions and Idioms in English: a Corpus Based Approach*. Oxford: Clarendon Press.
- Nattinger, James R. & Jeanette S. DeCarrico. 1992. *Lexical phrases and language teaching*. Oxford: Oxford University Press
- Pecina, Pavel (2009). *Lexical Association Measures, Collocation Extraction*. Institute of Formal and Applied Linguistics, Charles University, Prague.
- Stock, O. (1989). *Parsing with flexibility, dynamic strategies, and idioms in mind*, *Computational Linguistics*, 15.1.
- Williams, B. 2003. *Collocation with Advanced Levels 1 – not entirely proper/ appropriate/good?* <http://www.esol.britishcouncil.org>. (4 February 2013). [38] Wright, J. (1999). *Idioms organizer*, UK: Thomson Heinle.

## Dictionaries

- Longman dictionary of language Teaching and Applied Linguistics* (1992), by Jack Richards. & Richard Schmidt. Longman: Pearson Education
- Oxford Advanced Learner's Dictionary of Current English*. (1995) Oxford: Oxford University Press.
- Oxford Dictionary of Current Idiomatic English Vol. 2 – Phrasal Verbs*. 1983. Oxford: Oxford University Press.

## Sources

- Adib, A. (2010). *Persian Language and Literature*, زبان و ادبیات فارسی, [zaban va adabiate farsi]. Retrieved February 19, 2013, from <http://aryaadib.blogfa.com/cat-30.aspx>.
- Fararoy, J. (2009). [farhange Teyfi], فرهنگ طیفی, *Teyfi dictionary* (p. 747). Tehran: Hermes publication. Retrieved <http://www.scict.ir/portal/Home/Default.aspx?CategoryID=a88bd56-4ebd-4c32-80d0-a8406cd1655d>. Date Accessed: 23 February 2013
- Najafi, Abolhassan. (2000). [farhang farsi amiane], فرهنگ فارسی عامیانه, *Persian Slang Dictionary*. (First, P.700). Tehran: Nilofar. ISBN: 964-448-117-8. Date Accessed: 17 October 2013.
- Rahiminia, Mostafa. (2008). [zarbolmasalhaye farsi va estelahate amiane] و ضرب المثلهای فارسی و اصطلاحات عامیانه, *Persian Proverbs and Idiomatic Expressions*. (Sixth, P.200). Tehran: Golpa. ISBN: 978-964-7938-30-3. Date Accessed: 10 October 2013
- s/n. (2011). [fohshha va estelahate farsi be engelisi], فحش ها و اصطلاحا فارسی به انگلیسی, *Farsi to English Swear Words and Terminology*. Retrieved February 19, 2013, from <http://dellux.blogfa.com/post-21.aspx>.
- s/n. (2012). [estelahat, ebarate amiane, zarbolmasalha, jomalate por manai engelisi va afaale ebarati ba tajomei farsi] اصطلاحات, عبارات عامیانه, ضرب المثلهای, جملات پرمعناى انگلیسی و افعال ebarati ba tajomei farsi] *Idioms, Slang and Proverbs, English meaningful sentences and Verbs* (p. 79). Retrieved from. <http://www.kardoonline.com>. Date Accessed: 26 February 2013
- s/n. (2013). [estelahate amiane va rayej be zaban Farsi va engelisi], اصطلاحات عامیانه و رایج به زبان فارسی و انگلیسی, *Popular Idiomatic Expression in Persian and English*. Retrieved February 19, 2013, from <http://hedionline.com/index.php/learn-english/phrasal-verb-idiom-slang/142-2012-11-10-19-45-33>.
- s/n. (2013). [farhange estelahate amiane], فرهنگ اصطلاحات عامیانه, *Expressions Dictionary*. Retrieved February 19, 2013, from <http://www.sootak.ir>.
- Samaee, Seyed Mehdi. (2003). [farhang estelahate zabane makhfi], فرهنگ اصطلاحات زبان مخفی, *Hidden Language Glossary*. (D. A. Sadeghi, Ed.) (First., p. 101). Tehran: Ghazal. Retrieved from <http://forum.p30world.com/showthread.php?t=559720>. Date Accessed: 25 February 2013

## Other sources

- Khorasani, R. (1929). [farhange vajehaye amine (dar dorei Qhajar)], فرهنگ واژه های عامیانه در دوره قاجار, *Expression Glossary* (in Qajar period). (Nineteenth, P. 38). Tehran: [Farhangestan zaban va adabiate Farsi] فرهنگستان زبان و ادب فارسی, Academy of Persian Language and Literature. Retrieved from. <http://www.persianacademy.ir>. Date Accessed: February 25, 2013

- s/n. (2010). [entesharate jangal], انتشارات جنگل, *Jungle Publication*. Retrieved from. <http://www.junglepub.org>. Date Accessed: February 25, 2013
- s/n. (2011). [estelahat va tabirat], اصطلاحات و تعبيرات, *Expressions and interpretations*. Retrieved from. <http://www.englishpro.ir>. Date Accessed: February 25, 2013
- Arch , Sam. (2011). [estelahate amiane be zabane farsi va manie anha], اصطلاحات عاميانه زبان فارسی, و معانی آنها, *Persian slang terms and their meanings*. Retrieved from. <http://www.noandishaan.com/forums/thread40805.html>. Date Accessed: February 25, 2013

## Appendices

### Appendix A. Database

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	آب	<E>	آوردن	<E>	-	+	آب آوردن	ab avardana n	water bring	bringing water	being sick	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	آب	<E>	دادن	<E>	-	+	آب پس دادن	ab pas dadanan	water give back	giving back the water	being generous-being unreliable	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	+	آب	<E>	رفتن	<E>	-	+	آب رفتن	ab raftanan	water go	going water	became short-became low	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	آتش	<E>	کردن	<E>	-	+	آتش کردن	atash kardan	fire do	making fire	shooting	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	آتش	<E>	گرفتن	<E>	-	+	آتش گرفتن	atash gereftan	fire take	taking fire	get angry	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	+	آخ	<E>	گفتن	<E>	-	+	آخ گفتن	akh goftan	ouch say	saying ouch	not being strong	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	آزاد	<E>	ساختن	<E>	-	+	آزاد ساختن	azadanan sakhtan	free make	making free	returns	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	+	آشپز	<E>	گرفتن	<E>	-	+	آشپز گرفتن	ashian gereftan	nest get	getting nest	settle	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	آفتاب	<E>	چریدن	<E>	-	-	آفتاب چریدن	aftab charidan	sun pasture	pasture sun	having fun	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	آفتاب	<E>	خوردن	<E>	-	+	آفتاب خوردن	aftab khordan	sun eat	eating sun	being in a difficult situation	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	باب	<E>	کردن	<E>	-	-	باب کردن	bab kardan	door do	doing door	making something prevalent	<a href="http://www.scict.ir">http://www.scict.ir</a>
-	+	باد	<E>	خوردن	<E>	-	+	باد خوردن	bad khordan	wind eat	eating wind	something which is forgotten	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	باد	<E>	کردن	<E>	-	-	باد در کردن	bad darkardana n	wind drive out	driving the wind out	puke	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	بار	<E>	آوردن	<E>	-	+	بار آوردن	bar avardana n	load bring	bringing load	educate	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	باران	<E>	دیدن	<E>	-	+	باران دیدن	baran didan	rain see	seeing rain	impassive	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	بازی	<E>	دادن	<E>	-	+	بازی دادن	bazi dadan	play give	giving play	cheat	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	بازی	<E>	خوردن	<E>	-	-	بازی خوردن	bazi khordan	play eat	eating play	be deceived	<a href="http://www.scict.ir">http://www.scict.ir</a>

Persian idioms

N0=:Hum	N0=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
-	+	بالا	<E>	گرفتن	<E>	-	-	بالا گرفتن	bala gereftan	high get	getting high	intensify	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	بد	<E>	آوردن	<E>	-	+	بد آوردن	bad avardan	bad bring	bringing bad	being unlucky	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	بد	<E>	دیدن	<E>	-	+	بد دیدن	bad didan	bad see	seeing bad	being in a bad situation	<a href="http://www.scict.ir">http://www.scict.ir</a>
-	+	برق	<E>	افتادن	<E>	-	-	برق افتادن	bargh oftadan	electricity fall	falling electricity	being too shiny and clean	<a href="http://www.scict.ir">http://www.scict.ir</a>
-	+	بند	<E>	آمدن	<E>	-	-	بند آمدن	band amadan	dam come	coming dam	stopped	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	پيله	<E>	کردن	<E>	-	+	پيله کردن	pile kardan	Cocoon do	doing cocoon	bothering somebody by persistence	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	پز	<E>	دادن	<E>	-	-	پز دادن	poz dadan	posture give	giving posture	show off	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	تاب	<E>	آوردن	<E>	-	-	تاب آوردن	tab avardan	swing bring	bringing swing	resist	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	تاب	<E>	خوردن	<E>	-	-	تاب خوردن	tab khordan	swing eat	eating swing	going out to have fun	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	تابلو	<E>	کردن	<E>	-	-	تابلو کردن	tablo kardan	board do	doing board	revealing something	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	تپق	<E>	زدن	<E>	-	-	تپق زدن	topogh zadan	fetlock hit	hitting fetlock	having problem in talking	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	تندی	<E>	کردن	<E>	-	-	تندی کردن	tondi kardan	speed do	doing speed	being angry	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	ترش	<E>	کردن	<E>	-	+	ترش کردن	torsh kardan	sour do	doing sour	becoming angry	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	جا	<E>	زدن	<E>	-	-	جا زدن	ja zadan	place hit	hitting place	fearing to do something	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	جفت	<E>	کردن	<E>	-	+	جفت کردن	joft kardan	pair do	doing pair	fearing	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	خر	<E>	زدن	<E>	-	+	خر زدن	khar zadan	donkey hit	hitting donkey	studing too much	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	دبه	<E>	کردن	<E>	-	-	دبه کردن	dabe kardan	gallon do	doing gallon	renege	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	قر	<E>	دادن	<E>	-	-	قر دادن	gher dadan	concave give	giving concave	dancing	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	قر	<E>	آمدن	<E>	-	-	قر آمدن	gher amadan	dance come	coming dance	pretexting	<a href="http://www.scict.ir">http://www.scict.ir</a>

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	کام	<E>	گرفتن	<E>	-	-	کام گرفتن	kam gereftan	palate take	taking palate	enjoying something	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	کف	<E>	کردن	<E>	-	-	کف کردن	kam gereftan	foam do	doing foam	get excited	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	کم	<E>	آوردن	<E>	-	+	کم آوردن	kam avardan	low bring	brining low	being exhausted, fear	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	کم	<E>	داشتن	<E>	-	+	کم داشتن	kam dashtan	low have	having low	being idiot	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	گاف	<E>	دادن	<E>	-	-	گاف دادن	gaf dadan	gaf give	giving gaf	showing something up	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	مُقر	<E>	آمدن	<E>	-	-	مُقر آمدن	moghor amadan	talk come	coming talk	reveal The secret	<a href="http://www.scict.ir">http://www.scict.ir</a>
+	-	مخ	<E>	خوردن	<E>	-	-	مخ خوردن	mokh khordan	brain eat	eating brain	talking a lot	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	گیر	<E>	دادن	<E>	-	-	گیر دادن	gir dadan	snag give	giving snag	insisting	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	لُفت	<E>	دادن	<E>	+	-	لُفت دادن	left dadan	lengthen out give	giving lengthen out	moving like snail	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	کرم	<E>	ریختن	<E>	-	+	کرم ریختن	kerm rikhtan	worm drope	droping worms	teasing	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	کور	<E>	خواندن	<E>	-	-	کور خواندن	kor khand	blind read	reading blind	making mistake	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	کوتاه	<E>	آمدن	<E>	-	-	کوتاه آمدن	kotah amadan	short come	coming short	forget about it	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	گند	<E>	زدن	<E>	-	-	گند زدن	gand zadan	stench hit	hitting stench	sabotage	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	جوش	<E>	آوردن	<E>	-	-	جوش آوردن	josh avardan	boil bring	bringing boil	becoming angry	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	پس	<E>	رفتن	<E>	-	+	پس رفتن	pas gereftan	back go	going back	regression	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	جان	<E>	کندن	<E>	-	-	جان کندن	jan kandan	life dig	diging life	working too much	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	فخر	<E>	فروختن	<E>	-	-	فخر فروختن	fakhr frokhtan	glory sell	selling glory	show off	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	پز	<E>	دادن	<E>	-	-	پز دادن	poz dadan	posture give	giving posture	show off	<a href="http://www.kardano.com">http://www.kardano.com</a>
-	+	خاک	<E>	گرفتن	<E>	-	+	خاک گرفتن	khak gereftan	dust catch	catching dust	getting Old	<a href="http://www.kardano.com">http://www.kardano.com</a>
+	-	دل	<E>	ریختن	<E>	-	-	دل ریختن	del rikhtan	heart shed	shedding heart	fear	<a href="http://www.kardano.com">http://www.kardano.com</a>

Persian idioms

N0=:Hum	N0=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	+	آب	<E>	رفتن	<E>	-	+	آب رفتن	ab raftan	water go	going water	becoming thin	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	دم	<E>	زدن	<E>	-	-	دم زدن	dam zadan	breath hit	hitting breath	talking	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	خیت	<E>	کاشتن	<E>	-	-	خیت کاشتن	khit kashtan	thread plant	planting thread	sabotage	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	مزه	<E>	انداختن	<E>	-	-	مزه انداختن	maze andakhtan	taste throw	throwing taste	kidding	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	مزه	<E>	ریختن	<E>	-	-	مزه ریختن	maze rikhtan	taste shed	shedding taste	kidding	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	ناز	<E>	کردن	<E>	-	-	ناز کردن	naz kardan	Coquetry do	doing coquetry	play hard to get something	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	یکدستی	<E>	خوردن	<E>	-	+	یکدستی خوردن	yedasti khordan	handheld eat	eating handheld	be deceived	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	هرز	<E>	رفتن	<E>	-	-	هرز رفتن	harz raftan	wasted go	going wasted	footloose being	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	پوست	<E>	انداختن	<E>	-	+	پوست انداختن	post andakhtan	skin throw	throwing skin	have a job to do	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	کیسه	<E>	دوختن	<E>	-	+	کیسه دوختن	kise dokhtan	bag sew	sewing bag	have an eye on something	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	زور	<E>	زدن	<E>	-	-	زور زدن	zor zadan	force hit	hitting force	trying to do something	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	تن	<E>	خاریدن	<E>	-	+	تن خاریدن	tan kharidan	body itch	itching body	eager for the fray	<a href="http://www.kardanoonline.com">http://www.kardanoonline.com</a>
+	-	آمار	<E>	دادن	<E>	+	+	آمار دادن	amar dadan	Statistics give	giving statistics	sending a positive waves to make relation with opposite sex	<a href="http://forum.p30world.com/showthread.php?t=559720">http://forum.p30world.com/showthread.php?t=559720</a>
+	-	آمار	<E>	گرفتن	<E>	-	+	آمار گرفتن	amar gereftan	statistics take	taking statistics	trying to achieve some info	<a href="http://forum.p30world.com/showthread.php?t=559721">http://forum.p30world.com/showthread.php?t=559721</a>
+	-	آنتن	<E>	دادن	<E>	-	+	آنتن دادن	anten dadan	Antenna give	giving antenna	sending a positive waves to make relation with opposite sex	<a href="http://forum.p30world.com/showthread.php?t=559722">http://forum.p30world.com/showthread.php?t=559722</a>
+	-	اتو	<E>	زدن	<E>	-	-	اتو زدن	oto zadan	car hit	hitting car	getting taxi without paying	<a href="http://forum.p30world.com/show">http://forum.p30world.com/show</a>

Persian idioms

N0=:Hum	N0=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
													wthread.php?t=559723
+	-	باد	<E>	کردن	<E>	-	+	باد کردن	bad kardan	wind do	doing wind	getting pregnant	http://forum.p30world.com/showthread.php?t=559724
+	-	ترمز	<E>	بریدن	<E>	-	+	ترمز بریدن	tormoz boridan	Brake cut	cutting brake	becoming angry	http://forum.p30world.com/showthread.php?t=559725
+	-	تریپ	<E>	برداشتن	<E>	-	-	تریپ برداشتن	tirip bardashtan	trip take	taking trips	show off	http://forum.p30world.com/showthread.php?t=559726
+	-	قیافه	<E>	گرفتن	<E>	-	-	قیافه گرفتن	giafe gereftan	face take	taking face	show off	http://forum.p30world.com/showthread.php?t=559727
+	-	تریپ	<E>	زدن	<E>	-	-	تریپ زدن	tartib dadan	trip hit	hitting trip	wearing fashionable clothes to be handsome	http://forum.p30world.com/showthread.php?t=559728
+	-	تگری	<E>	زدن	<E>	-	-	تگری زدن	tagari zadan	cold hit	hitting cold	puking	http://forum.p30world.com/showthread.php?t=559729
+	-	چت	<E>	زدن	<E>	-	-	چت زدن	chet zadan	hazy hit	hitting hazy	bete noire being	http://forum.p30world.com/showthread.php?t=559730
+	-	چت	<E>	کردن	<E>	-	-	چت کردن	chet kardan	hazy do	doing hazy	having illusion	http://forum.p30world.com/showthread.php?t=559731
+	+	حال	<E>	دادن	<E>	-	-	حال دادن	hal dadan	mood give	giving mood	causing a passion and pleasure	http://forum.p30world.com/showthread.php?t=559732
+	-	خط	<E>	دادن	<E>	-	-	خط دادن	khat dadan	line give	giving line	sending a positive waves to make	http://forum.p30world.com/showthread.php

Persian idioms

N0=:Hum	N0=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
												relation with opposite sex	p?t=559733
+	-	خط	<E>	زدن	<E>	-	+	خط زدن	khat zadan	line hit	hitting line	having fun	http://forum.p30world.com/showthread.php?t=559734
+	-	راه	<E>	دادن	<E>	-	-	راه دادن	rah dadan	way give	giving way	sending a positive waves to make relation with opposite sex	http://forum.p30world.com/showthread.php?t=559735
+	-	رکب	<E>	خوردن	<E>	-	-	رکب خوردن	rakab khordan	trick eat	eating trick	be deceived	http://forum.p30world.com/showthread.php?t=559736
+	-	سوتی	<E>	دادن	<E>	-	-	سوتی دادن	soti dadan	Goofs give	giving goofs	telling something that cause revealing the secret	http://forum.p30world.com/showthread.php?t=559737
+	-	سه	<E>	کردن	<E>	-	-	سه کردن	se kardan	three do	doing three	telling something that cause revealing the secret	http://forum.p30world.com/showthread.php?t=559738
+	-	شیرین ن	<E>	زدن	<E>	-	-	شیرین زدن	shirin zadan	sweet hit	hitting sweet	looks stupid	http://forum.p30world.com/showthread.php?t=559740
+	-	فک	<E>	زدن	<E>	-	-	فک زدن	fak zadan	jaw hit	hitting jaw	talking too much	http://forum.p30world.com/showthread.php?t=559741
+	-	فر	<E>	زدن	<E>	-	-	فر زدن	fer zadan	curl hit	hitting curl	secret escape	http://forum.p30world.com/showthread.php?t=559742
+	-	فر	<E>	خوردن	<E>	-	-	فر خوردن	fer khordan	curl eat	eating curl	becoming angry	http://forum.p30world.com/showthread.php?t=559743

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	قاط	<E>	زدن	<E>	-	-	قاط زدن	ghat zadan	mix hit	hitting mix	becoming angry	<a href="http://forum.p30world.com/showthread.php?t=559744">http://forum.p30world.com/showthread.php?t=559744</a>
+	+	قاطی	<E>	کردن	<E>	-	+	قاطی کردن	ghati kardan	mix do	doing mix	becoming angry	<a href="http://forum.p30world.com/showthread.php?t=559745">http://forum.p30world.com/showthread.php?t=559745</a>
+	-	کف	<E>	ریختن	<E>	-	+	کف ریختن	kaf rikhtan	foam shed	shedding foam	becoming angry	<a href="http://forum.p30world.com/showthread.php?t=559746">http://forum.p30world.com/showthread.php?t=559746</a>
+	-	کلاس	<E>	گذاشتن	<E>	-	+	کلاس گذاشتن	kelas gozashtan	class put	putting class	Mimic the behavior of educated people	<a href="http://forum.p30world.com/showthread.php?t=559747">http://forum.p30world.com/showthread.php?t=559747</a>
+	-	کلید	<E>	کردن	<E>	-	-	کلید کردن	kelid kardan	key do	doing key	bete noire being	<a href="http://forum.p30world.com/showthread.php?t=559748">http://forum.p30world.com/showthread.php?t=559748</a>
+	-	مخ	<E>	زدن	<E>	-	-	مخ زدن	mokh zadan	brain hit	hitting brain	persuade somebody by talking	<a href="http://forum.p30world.com/showthread.php?t=559749">http://forum.p30world.com/showthread.php?t=559749</a>
+	-	نخ	<E>	دادن	<E>	-	+	نخ دادن	nakh dadan	thread give	giving thread	sending a positive waves to make relation with opposite sex	<a href="http://forum.p30world.com/showthread.php?t=559750">http://forum.p30world.com/showthread.php?t=559750</a>
+	-	امپر	<E>	چسباندن	<E>	-	-	امپر چسباندن	amper chasbandan	ampere past	pasting ampere	becoming angry	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	ارایه	<E>	دادن	<E>	-	-	ارایه دادن	araye dadan	array give	giving array	messing something up	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	پا	<E>	دادن	<E>	-	+	پا دادن	pa dadan	foot give	giving foot	sending a positive waves to make relation with opposite sex	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	تخم	<E>	کردن	<E>	-	+	تخم کردن	tokhm kardan	testicel do	doing testicel	greatly fear	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	تیکه	<E>	انداختن	<E>	-	+	تیکه انداختن	tike andakhtan	piece drop	Dropping piece	saying ironic words to bother the others	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	بندری	<E>	زدن	<E>	-	-	بندری زدن	bandari zadan	port hit	hitting port	someone who seems boob	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	پاچه	<E>	خارانند	<E>	-	+	پاچه خارانند	pache kharand	leg scratch	scratching leg	flatter someone	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	تیغی	<E>	زدن	<E>	-	-	تیغی زدن	tighi zadan	blade hit	hitting blade	betting	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	ردیف	<E>	کردن	<E>	-	+	ردیف کردن	radif kardan	row do	doing row	making everything alright	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	زر	<E>	زدن	<E>	-	-	زر زدن	zer zadan	blather hit	hitting blather	twaddle	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	قات	<E>	زدن	<E>	-	-	قات زدن	ghat zadan	mix hit	hitting mix	becoming angry	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	کره	<E>	کردن	<E>	-	-	کره کردن	kare kardan	butter do	doing butter	eating too much	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	گوجه	<E>	زدن	<E>	-	+	گوجه زدن	goje zadan	tomato hit	hitting tomato	puking	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	لاو	<E>	انداختن	<E>	-	-	لاو انداختن	love andakhtan	love drop	dropping love	being in love	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-	لاو	<E>	ترکاندن	<E>	-	-	لاو ترکاندن	love terekandan	love blast	blasting love	being in love	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>

Persian idioms

	N0=:Hum	N0=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-		لاي	<E>	کشيدن	<E>	-	+	لاي کشيدن	lai keshidan	pad pull	pulling pad	being a hypocrite	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-		نبشي	<E>	دادن	<E>	-	+	نبشي دادن	nabshi dadan	cornerstone give	giving Cornerstone	telling something that cause revealing the secret	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-		ضايح	<E>	کردن	<E>	-	-	ضايح کردن	zaye kardan	damage do	doing damage	messing something up	<a href="http://aryaadib.blogfa.com/cat-30.aspx">http://aryaadib.blogfa.com/cat-30.aspx</a>
+	-		آشنايي	<E>	دادن	<E>	-	-	آشنايي دادن	ashenai dadan	familiarity give	giving familiarity	in the presence of others showing that you know a special person	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		افاده	<E>	فروختن	<E>	-	-	افاده فروختن	efade frokhtan	pride sell	selling pride	show off	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		افلاس	<E>	افتادن	<E>	-	-	افلاس افتادن	eflas oftadan	bankruptcy fall	falling bankruptcy	getting caught in the poverty	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		بار	<E>	انداختن	<E>	-	+	بار انداختن	bar andakhtan	load drop	dropping load	settle somewhere	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		بو	<E>	بردن	<E>	-	-	بو بردن	bo bordan	odor take	taking odor	understand something based on evidence	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		پر	<E>	آوردن	<E>	-	+	پر در آوردن	par daravardan	cilium evolve	evolve cilium	becoming very happy	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		پرسه	<E>	زدن	<E>	-	-	پرسه زدن	parse zadan	prowl hit	hitting prowl	walking without a goal	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		پف	<E>	کردن	<E>	-	+	پف کردن	pof kardan	puff do	doing puff	swelling because of sickness	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		ترش	<E>	کردن	<E>	-	+	ترش کردن	torsh kardan	sour do	doing sour	becoming angry	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		جا	<E>	خوردن	<E>	-	-	جا خوردن	ja khordan	place eat	eating place	becoming surprised by seeing or hearing something weird	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-		چشم	<E>	انداختن	<E>	-	-	چشم انداختن	cheshm andakhtan	eye Drop	dropping eye	looking casually	<a href="http://www.sootak.ir">http://www.sootak.ir</a>

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	چونه	<E>	زدن	<E>	-	-	چونه زدن	chone zadan	chin hit	hitting chin	asking for discount	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-	حرف	<E>	آوردن	<E>	-	+	حرف در آوردن	harf daravardan	word make	making word	rumors making	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-	خشك	<E>	زدن	<E>	-	-	خشك زدن	khoshk zadan	dry hit	hitting dry	stay flabbergasted	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-	خيز	<E>	برداشتن	<E>	-	-	خيز برداشتن	khiz bardashtan	jump remove	removing jump	ready to attack someone or something	<a href="http://www.sootak.ir">http://www.sootak.ir</a>
+	-	ابغور	<E>	گرفتن	<E>	-	+	ابغوره گرفتن	abghore gereftan	verjuice get	getting verjuice	cry	Rahiminia (2008)
+	-	استخوان ان	<E>	کردن	<E>	-	+	استخوان خرد کردن	ostokhan khordan kardan	bone crushing	crushing bone	suffer in learning something	Rahiminia (2008)
+	-	باج سبيل	<E>	دادن	<E>	+	-	باج سبيل دادن	baj sebil dadan	blackmail give	giving blackmail	taking money by force	Rahiminia (2008)
+	-	نیش	<E>	زدن	<E>	-	+	نیش زدن	nish zadan	sting hit	hitting sting	squibbing	Rahiminia (2008)
+	-	بازار گرمی	<E>	کردن	<E>	-	-	بازار گرمی کردن	bazargar mi kardan	market warmth do	doing warmth market	high praise of seller, for its goods	Rahiminia (2008)
+	-	بازی	<E>	آوردن	<E>	-	-	بازی در آوردن	bazi daravardan	game evolve	evolve game	evade	Rahiminia (2008)
+	-	بامبول	<E>	آوردن	<E>	-	-	بامبول در آوردن	bambol daravardan	flam evolve	evolve flam	flam	Rahiminia (2008)
+	-	بامبول	<E>	زدن	<E>	-	-	بامبول زدن	bambol zadan	flam hit	hit flame	flam	Rahiminia (2008)
+	-	بل	<E>	گرفتن	<E>	-	-	بل گرفتن	bol gereftan	ball take	taking ball	best use of the opportunity	Rahiminia (2008)
+	-	زمین	<E>	خوردن	<E>	-	-	زمین خوردن	zamin khordan	ground eat	eating ground	a person who lost everything	Rahiminia (2008)
+	-	بیخ	<E>	کردن	<E>	-	-	بیخ پیدا کردن	bikh peyda kardan	butt find	finding butt	anything that become complicated	Rahiminia (2008)
+	-	پاپوش	<E>	دوختن	<E>	+	-	پاپوش دوختن	paposh dokhtan	Shoes stitch	stitching shoes	causing Conspiracy	Rahiminia (2008)
+	-	پرپر	<E>	زدن	<E>	-	+	پرپر زدن	par par zadan	Pitter-patter hit	hitting Pitter-patter	being Impatient	Rahiminia (2008)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	پرده پوشی	<E>	کردن	<E>	-	-	پرده پوشی کردن	pardeposhi kardan	Curtain Sheath do	doing Curtain Sheath	hiding somebody's fault	Rahiminia (2008)
+	-	پس	<E>	افتادن	<E>	-	-	پس افتادن	pas oftadan	back fall	falling back	fainting	Rahiminia (2008)
+	-	پس	<E>	شاشیدن	<E>	-	-	پس شاشیدن	pas shashidan	back piss	pissing back	retrogress	Rahiminia (2008)
+	-	پنجر	<E>	کردن	<E>	-	+	پنجر کردن	panchar kardan	puncture do	doing puncture	being too tired	Rahiminia (2008)
+	-	پيله	<E>	کردن	<E>	-	+	پيله کردن	pile kardan	cocoon do	doing cocoon	insisting to get something	Rahiminia (2008)
+	-	تخم	<E>	گذاشتن	<E>	-	+	تخم گذاشتن	tokhm gozashtan	egg put	putting egg	doing wrong act	Rahiminia (2008)
+	-	تاوان	<E>	دادن	<E>	-	-	تاوان پس دادن	tavan pas dadan	compensation give back	giving back compensation	damage compensation	Rahiminia (2008)
+	-	تخت	<E>	خوابیدن	<E>	-	-	تخت خوابیدن	takht khabidan	couch sleep	sleeping couch	sleeping like a baby	Rahiminia (2008)
+	-	ترکمون	<E>	زدن	<E>	+	-	ترکمون زدن	terkamon zadan	stench hit	hitting stench	sabotage	Rahiminia (2008)
+	-	جر	<E>	زدن	<E>	-	-	جر زدن	jer zadan	cheat hit	hitting cheat	cheating	Rahiminia (2008)
+	-	جفتک	<E>	انداختن	<E>	-	+	جفتک انداختن	joftak andakhtan	capriole throw	throwing capriole	misbehaving	Rahiminia (2008)
+	-	جفتگ	<E>	گفتن	<E>	-	-	جفتگ گفتن	jafang goftan	fudge say	saying fudge	trash talk	Rahiminia (2008)
+	-	چیک	<E>	زدن	<E>	-	-	چیک زدن	jik zadan	chirp hit	hitting chirp	protest	Rahiminia (2008)
+	-	چیم	<E>	زدن	<E>	-	-	چیم زدن	jim zadan	scram hit	hitting scram	french leave	Rahiminia (2008)
+	-	چاپ	<E>	زدن	<E>	-	+	چاپ زدن	chap zadan	print hit	hitting print	telling lie	Rahiminia (2008)
+	-	چو	<E>	انداختن	<E>	+	-	چو انداختن	cho andakhtan	rumor throw	throwing rumor	broadcast a wrong news	Rahiminia (2008)
+	-	حاشیه	<E>	رفتن	<E>	-	-	حاشیه رفتن	hashie raftan	border go	going border	being away from the core of talk	Rahiminia (2008)
+	-	حال	<E>	آمدن	<E>	-	-	حال آمدن	hal amadan	mood come	coming mood	becoming healthy	Rahiminia (2008)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	ح ص	<E>	خوردن	<E>	-	-	حرص خوردن	hers khordan	greed eat	eating greed	being angry and anxious	Rahiminia (2008)
+	-	ح ص	<E>	زدن	<E>	-	-	حرص زدن	hers zadan	greed hit	hitting greed	trying to get more than your needs	Rahiminia (2008)
+	-	حرف	<E>	آوردن	<E>	-	-	حرف در آوردن	harf daravardan	word evolve	evolve word	gossiping	Rahiminia (2008)
+	-	حساب	<E>	دادن	<E>	-	+	حساب پس دادن	hesab pas dadan	account give back	give back account	responding to boss questions	Rahiminia (2008)
+	-	حلاجی	<E>	کردن	<E>	-	+	حلاجی کردن	halaji kardan	batting do	do batting	examine with high level of carefulness	Rahiminia (2008)
+	-	خرده شیشه	<E>	داشتن	<E>	-	+	خرده شیشه داشتن	khordane shishe dashtan	splinter have	having splinter	being wicked and tricky	Rahiminia (2008)
+	-	خون	<E>	کردن	<E>	-	-	خون کردن	khon kardan	blood do	doing blood	being a murderer	Rahiminia (2008)
+	-	درز	<E>	گرفتن	<E>	-	+	درز گرفتن	darz gerftan	seam take	taking seam	talking briefly	Rahiminia (2008)
+	-	دلخور	<E>	داشتن	<E>	-	-	دلخوری داشتن	delkhori dashtan	annoyance have	having annoyance	being dissatisfied	Rahiminia (2008)
+	-	دل دل	<E>	کردن	<E>	-	-	دل دل کردن	del del kardan	heart heart do	doing heart heart	vacillate	Rahiminia (2008)
+	-	دم	<E>	آوردن	<E>	-	+	دم در آوردن	dom daravardan	tail evolve	evolve tail	becoming barefaced	Rahiminia (2008)
+	-	ذوق	<E>	کردن	<E>	-	-	ذوق کردن	zogh kardan	taste do	doing taste	becoming happy	Rahiminia (2008)
+	-	رکاب	<E>	دادن	<E>	-	+	رکاب دادن	rekab dadan	pedal give	giving pedal	give up	Rahiminia (2008)
+	-	کف	<E>	رفتن	<E>	-	-	کف رفتن	kaf raftan	floor go	going floor	stealing	Rahiminia (2008)
+	-	رو	<E>	انداختن	<E>	-	-	رو انداختن	ro andakhtan	face throw	throwing face	asking	Rahiminia (2008)
+	-	رو	<E>	داشتن	<E>	-	-	رو داشتن	ro dashtan	face have	having face	being barefaced	Rahiminia (2008)
+	-	سر	<E>	آوردن	<E>	-	+	سر در آوردن	sar daravardan	head evolve	evolve head	understand	Rahiminia (2008)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	سماق	<E>	مکین	<E>	-	+	سماق مکین	somagh mekidan	sumac suck	sucking sumac	wating for something	Rahiminia (2008)
+	-	سنگ	<E>	انداختن	<E>	+	+	سنگ انداختن	sang andakhtan	stone throw	throwing stone	causing problem for someone	Rahiminia (2008)
+	-	سوسه	<E>	آمدن	<E>	-	-	سوسه آمدن	sose amadan	weevil come	coming weevil	causing problem for someone	Rahiminia (2008)
+	-	صغرا کبرا	<E>	چیدن	<E>	-	-	صغرا کبرا چیدن	soghra kobra chidan	cobra soghra kobra	picking cobra soghra	not talking about major point of the topic	Rahiminia (2008)
+	-	غاز	<E>	چراندن	<E>	-	+	غاز چراندن	ghaz charandan	goose graz	grazing goos	being unemployed	Rahiminia (2008)
+	-	قاپ	<E>	زدن	<E>	-	-	قاپ زدن	ghap zadan	rogue hit	hitting rogue	sudden stealing	Rahiminia (2008)
+	-	قار قار	<E>	کردن	<E>	-	+	قار قار کردن	ghar ghar kardan	caw caw do	doing caw caw	being talkative	Rahiminia (2008)
+	-	قپی	<E>	آمدن	<E>	-	-	قپی آمدن	ghopi amadan	boast come	coming boast	jactitation	Rahiminia (2008)
+	-	قمپز	<E>	کردن	<E>	-	-	قمپز در کردن	gompoz darkardan	bluff shoot	shooting bluff	a person who is bluffing	Rahiminia (2008)
+	-	لب	<E>	چیدن	<E>	-	-	لب برچیدن	lab barchidan	lip remove	removing lip	being ready to cry	Rahiminia (2008)
+	-	لب	<E>	گزیدن	<E>	-	+	لب گزیدن	lab gazidan	lip bit	biting lip	feel sorry	Rahiminia (2008)
+	-	لک لک	<E>	کردن	<E>	-	-	لک لک کردن	lek lek kardan	demoiselle do	doing demoiselle	doing something very slowly	Rahiminia (2008)
+	-	لنگ	<E>	انداختن	<E>	-	+	لنگ انداختن	long andakhtan	textile throw	throwing textile	give up	Rahiminia (2008)
+	-	لیچار	<E>	گفتن	<E>	+	-	لیچار گفتن	lychar goftan	lychar say	lychar saying	trash talk	Rahiminia (2008)
+	-	نم	<E>	دادن	<E>	-	+	نم پس دادن	nam pas nadadan	humidity give back	giving back humidity	being generous	Najafi (2000)
-	+	آب	<E>	کردن	<E>	-	+	آب باز کردن	ab baz kardan	water open	opening water	happening something wrong	Najafi (2000)
-	+	آب	<E>	برداشتن	<E>	-	+	آب برداشتن	ab bardashtan	water picking up	pick up water	being suspicious	Najafi (2000)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	آب	<E>	دادن	<E>	-	+	آب دادن	ab dadan	water give	giving water	reaching to orgasm	Najafi (2000)
-	+	آب	<E>	دیدن	<E>	-	+	آب دیدن	ab didan	water see	seeing water	damage	Najafi (2000)
+	-	آب	<E>	گوزیدن	<E>	-	-	آب گوزیدن	ab gozidan	water fart	fart water	cancel	Najafi (2000)
+	-	آتش	<E>	سوزاندن	<E>	-	-	آتش سوزاندن	atash sozandan	fire Burne	burning fire	causing riot	Najafi (2000)
+	-	آشنا	<E>	آمدن	<E>	-	-	آشنا در آمدن	ashena daramadan	familiar eventuate	eventuate familiar	knowing somebody by sudden	Najafi (2000)
+	-	آلودگی	<E>	زدن	<E>	-	+	آلودگی بهم زدن	alodegi beham zadan	pollution disarrange	disarrange pollution	getting married	Najafi (2000)
+	-	ابرو	<E>	آمدن	<E>	-	-	ابرو آمدن	abro amadan	eyebrows come	coming eyebrows	showing coquetry with shaking the eyebrows	Najafi (2000)
+	-	لنگ	<E>	کردن	<E>	-	+	لنگ کردن	lang kardan	lame do	doing lame	temporary living somewhere	Najafi (2000)
+	-	اثر	<E>	زدن	<E>	-	-	اثر زدن	eter zadan	ether hit	hitting ether	saying negative words and being pessimistic	Najafi (2000)
+	-	اجابت	<E>	گرفتن	<E>	-	-	اجابت گرفتن	ejabat gereftan	compliance get	getting compliance	the wishes which come true by pray	Najafi (2000)
+	-	خدمت	<E>	رفتن	<E>	-	-	خدمت رفتن	khdmata raftan	service go	going service	going to military service	Najafi (2000)
+	-	احلیل	<E>	زدن	<E>	-	-	احلیل زدن	ahlil zadan	penis hit	hitting penis	cheat somebody	Najafi (2000)
+	-	احلیل	<E>	خوردن	<E>	-	+	احلیل خوردن	ahlil khordan	penis eat	eating penis	somebody cheat you	Najafi (2000)
+	-	از	<E>	آمدن	<E>	-	-	از آمدن	ez amadan	inability come	coming inability	showing disability	Najafi (2000)
+	-	استخوان	<E>	ترکاندن	<E>	-	+	استخوان ترکاندن	ostokhan terekandan	bone blast	blasting bone	growing up very soon and getting weight	Najafi (2000)
+	-	استخوان	<E>	داشتن	<E>	-	+	استخوان داشتن	ostokhan dashtan	bone have	having bone	being experienced	Najafi (2000)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	اسم	<E>	آوردن	<E>	-	-	اسم در آوردن	esm daravardan	name evolve	evolve name	becoming famous	Najafi (2000)
+	-	اطوار	<E>	ریختن	<E>	-	-	اطوار ریختن	atvar rikhtan	modes shed	Shedding modes	showing coquetry with shaking hands	Najafi (2000)
+	-	اطوار	<E>	آوردن	<E>	-	-	اطوار در آوردن	atvar dar anvardan	modes evolve	evolve modes	showing coquetry with shaking hands	Najafi (2000)
+	-	اطوار	<E>	آمدن	<E>	-	-	اطوار آمدن	atvar amadan	modes come	come modes	showing coquetry with shaking hands	Najafi (2000)
-	+	افاقه	<E>	کردن	<E>	-	-	افاقه کردن	efaghe kardan	usefulness do	doig usefulness	having good effect	Najafi (2000)
+	-	تلب	<E>	شدن	<E>	-	-	تلب شدن	telep shodan	Plop become	become plop	going to somebody's house without invitation	Najafi (2000)
+	-	افه	<E>	آمدن	<E>	-	-	افه آمدن	efe amadan	efe come	coming efe	speaking and behaving odd to attract others	Najafi (2000)
+	-	امتحان	<E>	دادن	<E>	-	+	امتحان پس دادن	emtehan pas dadan	exam give back	giving back exam	showing the capability and incapability in specific field	Najafi (2000)
+	-	بال	<E>	آوردن	<E>	-	+	بال در آوردن	bal daravardan	wing evolve	evolve wing	becoming very happy	Najafi (2000)
+	-	گند	<E>	آوردن	<E>	-	-	گند بالا آوردن	gand bala avordan	stench raise	raising stench	any undermining of a cause, plan, or effort.	Najafi (2000)
+	-	خیت	<E>	آوردن	<E>	-	-	خیت بالا آوردن	khit bala avardan	thread raise	raising thread	any undermining of a cause, plan, or effort.	Najafi (2000)
+	-	حساب	<E>	آوردن	<E>	-	-	حساب بالا آوردن	hesab bala avardan	account raise	raising account	being debtor	Najafi (2000)

Persian idioms

N0=:Hum	N0=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	منبر	<E>	رفتن	<E>	-	-	منبر رفتن	menbar raftan	tribune go	going tribune	talking to much and giving advice to people	Najafi (2000)
+	-	بچگی	<E>	کردن	<E>	-	-	بچگی کردن	bachegi kardan	childhood do	doing childhood	doing something wrong and foolish because of inexperience	Najafi (2000)
+	-	بچه	<E>	انداختن	<E>	-	+	بچه انداختن	bache andakhtan	child drop	dropping child	Induced termination of a pregnancy with destruction of the embryo or fetus	Najafi (2000)
+	-	بچه	<E>	کردن	<E>	-	+	بچه درست کردن	bache dorost kardan	child make	making child	reproduction	Najafi (2000)
+	-	بخار	<E>	داشتن	<E>	-	+	بخار داشتن	bokhar dashtan	steam have	having Steam	being competent and capable	Najafi (2000)
+	-	مو	<E>	برداشتن	<E>	-	+	مو برداشتن	mo bardashtan	hair pick up	picking up hair	fine crack	Najafi (2000)
+	-	فیوز	<E>	پراندن	<E>	-	-	فیوز پراندن	fioz parandan	fuse jump	jumping fuse	getting shocked	Najafi (2000)
+	-	برق	<E>	پراندن	<E>	-	-	برق پراندن	bargh parandan	electricity jump	jumping electricity	getting shocked	Najafi (2000)
+	-	برگ	<E>	انداختن	<E>	-	+	برگ انداختن	barg andakhtan	leaf drop	dropping leaf	exaggerating a story with useless description	Najafi (2000)
+	-	برگ	<E>	زدن	<E>	-	+	برگ زدن	barg zadan	leaf hit	hitting leaf	cheating	Najafi (2000)
+	-	برو برو	<E>	داشتن	<E>	-	-	برو برو داشتن	boro boro dashtan	honor have	having honor	being so famous	Najafi (2000)
-	+	درز	<E>	کردن	<E>	-	+	درز کردن	darz kardan	seam do	doing seam	make known something private or secret	Najafi (2000)
+	-	بز	<E>	آوردن	<E>	-	+	بز آوردن	boz avaradan	goat bring	bringing goat	being unlucky	Najafi (2000)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	بز	<E>	گرفتن	<E>	-	+	بز گرفتن	boz gereftan	goat get	getting goat	buying a desirable good with low price	Najafi (2000)
+	-	بلبلی	<E>	خواندن	<E>	-	+	بلبلی خواندن	bolboli khandan	nightingales sing	singing nightingales	talking unreasonably	Najafi (2000)
+	-	بمب	<E>	انداختن	<E>	-	+	بمب انداختن	bomb andakhtan	drop bomb	dropping bomb	eating a piece of opium	Najafi (2000)
+	-	بند	<E>	کردن	<E>	-	-	بند کردن	band kardan	string do	doing string	insisting to do something	Najafi (2000)
+	-	رگ	<E>	کردن	<E>	-	-	رگ کردن	rag kardan	vessel do	doing vessel	stimulate of sexual desire	Najafi (2000)
+	-	غوغا	<E>	کردن	<E>	-	-	غوغا کردن	ghogha kardan	melee do	doing melee	doing the best	Najafi (2000)
+	-	قیامت	<E>	کردن	<E>	-	-	قیامت کردن	ghiamat kardan	judgment do	doing judgment	doing the best	Najafi (2000)
+	-	کولاک	<E>	کردن	<E>	-	-	کولاک کردن	kolak kardan	blizzard do	doing blizzard	doing the best	Najafi (2000)
+	-	پا	<E>	انداختن	<E>	-	+	پا انداختن	pa andakhtan	foot drop	dropping foot	preparing the facilities to do something	Najafi (2000)
+	-	پا	<E>	آوردن	<E>	-	+	پا در آوردن	pa dar avardan	foot evolve	evolve foot	doing something without permission	Najafi (2000)
+	-	پارتی	<E>	داشتن	<E>	-	+	پارتی داشتن	party dashtan	party have	having party	having wealthy friends	Najafi (2000)
+	+	پدر	<E>	آوردن	<E>	-	-	پدر در آوردن	pedar dar avardan	father evolve	evolve father	irritate or torment persistently	Najafi (2000)
+	-	پر	<E>	کشیدن	<E>	-	+	پر کشیدن	par keshidan	cilium draw	drawing cilium	die	Najafi (2000)
+	-	پز	<E>	آمدن	<E>	-	-	پز آمدن	poz amadan	posture come	coming posture	show off	Najafi (2000)
+	-	تقاص	<E>	دادن	<E>	-	-	تقاص پس دادن	taghas pas dadan	penalty give back	giving back penalty	paying the penalty of sin	Najafi (2000)
+	-	پشک	<E>	انداختن	<E>	-	-	پشک انداختن	peshk andakhtan	cat drop	dropping cat	doing lottery by hand fingers	Najafi (2000)
+	-	پک	<E>	زدن	<E>	-	-	پک زدن	pok zadan	puff hit	hitting puff	smoking	Najafi (2000)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	لنگر	<E>	انداختن	<E>	-	+	لنگر انداختن	langar andakhtan	anchor drop	dropping anchor	staying somewhere for a long time	Najafi (2000)
+	-	پولتیک کا	<E>	زدن	<E>	-	-	پولتیک زدن	poletik zadan	politic hit	hitting politic	cheating	Najafi (2000)
+	-	پوست	<E>	کندن	<E>	-	+	پوست کندن	post kandan	skin cut out	cut out skin	causing a horrible pain	Najafi (2000)
+	-	پیچ	<E>	خوردن	<E>	-	+	پیچ خوردن	pich khordan	twist eat	eating twist	somebody who try to get rid of you	Najafi (2000)
+	-	پیشانی ی	<E>	داشتن	<E>	-	+	پیشانی داشتن	pishani dashtan	forehead have	having forehead	being lucky	Najafi (2000)
+	-	پیه	<E>	آوردن	<E>	-	-	پیه آوردن	pih avardan	fat bring	bringing fat	getting fat	Najafi (2000)
+	-	تاب	<E>	برداشتن	<E>	-	+	تاب برداشتن	tab bardashtan	twist pick up	pick up twist	being muggles	Najafi (2000)
+	-	تاپاله	<E>	چسباندن	<E>	-	+	تاپاله چسباندن	tapale chasbandan	shit paste	Pasting shit	kissing loudly	Najafi (2000)
+	-	دور	<E>	برداشتن	<E>	-	-	دور برداشتن	dor bardashtan	cycle pick up	pick up cycle	asking for battle	Najafi (2000)
+	-	شلتاق	<E>	کردن	<E>	-	-	شلتاق کردن	shaltagh kardan	plunder do	doing plunder	asking for battle	Najafi (2000)
+	-	میدان	<E>	گرفتن	<E>	-	+	میدان گرفتن	meydan gereftan	battlefield take	taking battlefield	asking for battle	Najafi (2000)
+	-	تخم	<E>	کاشتن	<E>	-	+	تخم کاشتن	tokhm kashtan	seed plant	planting seed	reproduction	Najafi (2000)
+	-	تر	<E>	زدن	<E>	+	-	تر زدن	ter zadan	shit hit	hitting shit	Sabotage	Najafi (2000)
+	-	بهانه	<E>	تراشیدن	<E>	-	-	بهانه تراشیدن	bahane tarashidan	excuse shave	shaving excuse	always have new excuse	Najafi (2000)
+	-	دشمن	<E>	تراشیدن	<E>	+	-	دشمن تراشیدن	doshman tarashidan	enemy shave	shaving enemy	a guy who makes people become his enemy	Najafi (2000)
+	-	خرج	<E>	تراشیدن	<E>	+	-	خرج تراشیدن	kharj tarashidan	cost shave	shaving cost	doing something that cause spend money	Najafi (2000)
+	-	لب	<E>	ترکوندن	<E>	-	-	لب ترکوندن	lab terekandan	blast lip	lip blasting	start talking	Najafi

Persian idioms

N0=:Hum	N0=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
													(2000)
+	-	تشر	<E>	زدن	<E>	-	-	تشر زدن	tashar zadan	threat hit	hitting threat	browbeat	Najafi (2000)
+	-	تک	<E>	افتادن	<E>	-	-	تک افتادن	tak oftadan	monad drop	dropping monad	being alone	Najafi (2000)
+	-	توپ	<E>	آمدن	<E>	-	+	توپ آمدن	top amadan	ball come	coming ball	threaten	Najafi (2000)
+	-	تیار	<E>	آوردن	<E>	-	-	تیار در آوردن	tiart dar avardan	farce evolve	evolve farce	doing farce	Najafi (2000)
+	-	تیغ	<E>	زدن	<E>	-	+	تیغ زدن	tigh zadan	blade hit	hitting blade	taking money by force or trick	Najafi (2000)
+	-	جا	<E>	افتادن	<E>	-	+	جا افتادن	ja oftadan	place drop	dropping place	get the habit to new place	Najafi (2000)
+	-	یکه	<E>	خوردن	<E>	-	-	یکه خوردن	yeke khordan	unit eat	eating unit	being shocked	Najafi (2000)
+	-	جان	<E>	گرفتن	<E>	-	-	جان گرفتن	jan gereftan	life get	getting life	getting strong	Najafi (2000)
+	-	جان	<E>	دادن	<E>	-	-	جان دادن	jan dadan	life give	giving life	die	Najafi (2000)
+	-	جان	<E>	داشتن	<E>	-	-	جان داشتن	jan dashtan	life have	having life	being alive	Najafi (2000)
+	-	جواب	<E>	دادن	<E>	-	+	جواب پس دادن	javab pas dadan	answer give back	give back answer	justifying the act which is done	Najafi (2000)
+	-	جواب	<E>	گرفتن	<E>	+	-	جواب گرفتن	javab gereftan	answer get	getting answer	reaching to ideal result	Najafi (2000)
+	-	جوش	<E>	خوردن	<E>	-	-	جوش خوردن	josh khordan	boil eat	eating boil	becoming angry	Najafi (2000)
+	-	جوش	<E>	زدن	<E>	-	+	جوش زدن	josh zadan	boil hit	hitting boil	be concerned	Najafi (2000)
+	-	جوهر	<E>	داشتن	<E>	-	+	جوهر داشتن	johar dashtan	ink have	having ink	being competent and talented	Najafi (2000)
+	-	جیز	<E>	رفتن	<E>	-	-	جیز رفتن	jiz raftan	lucre go	going lucre	stealing the worthless thing	Najafi (2000)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	چیز	<E>	خوردن	<E>	-	-	چیز خوردن	jiz khordan	lucre eat	eating lucre	lend money to someone and ask to pay back double	Najafi (2000)
+	-	نطق	<E>	کشیدن	<E>	-	+	نطق کشیدن	notogh keshidan	speech drawl	drawl speech	express strong objection	Najafi (2000)
+	-	چیک	<E>	نشستن	<E>	-	-	چیک نشستن	jik neshastan	frame sit	sitting frame	sitting cross-legged	Najafi (2000)
+	-	چانه	<E>	داشتن	<E>	-	+	چانه داشتن	chane dashtan	chin have	having Chin	being talkative	Najafi (2000)
+	-	چانه	<E>	انداختن	<E>	-	+	چانه انداختن	chane andakhtan	chin throw	throwing chin	cease living; become dead	Najafi (2000)
+	-	چپه	<E>	کردن	<E>	-	+	چپه کردن	chape kardan	reversal do	doing reversal	being so upset	Najafi (2000)
+	-	چرتکه که	<E>	انداختن	<E>	-	+	چرتکه انداختن	chortke andakhtan	abacus throw	throwing abacus	anyone who care about his own profit	Najafi (2000)
+	-	چوسی	<E>	آمدن	<E>	-	-	چوسی آمدن	chosi amadan	fart come	coming fart	telling lie	Najafi (2000)
+	-	چشم	<E>	خوردن	<E>	-	+	چشم خوردن	cheshm khordan	eye eat	eating eye	being exposed to the evil eye	Najafi (2000)
+	-	نظر	<E>	خوردن	<E>	-	-	نظر خوردن	nazar khordan	sight eat	eating sight	being exposed to the evil eye	Najafi (2000)
+	-	چشم	<E>	برگرداندن	<E>	-	+	چشم برگرداندن	cheshm bargardandan	eye return	returning eye	neglecting for a second	Najafi (2000)
+	-	چل	<E>	زدن	<E>	-	-	چل زدن	chel zadan	insane hit	hitting insane	sounds stupid	Najafi (2000)
+	-	چله	<E>	افتادن	<E>	-	-	چله افتادن	chele oftdan	bowstring drop	dropping bowstring	temporary infertility of women after marriage	Najafi (2000)
+	-	چله	<E>	انداختن	<E>	-	-	چله انداختن	chele andakhtan	bowstring throw	throwing bowstring	stopping the progress of something	Najafi (2000)
+	-	چله	<E>	بریدن	<E>	-	-	چله بریدن	chele boridan	bowstring cut	cutting bowstring	trying to stop unlucky which is happening	Najafi (2000)

Persian idioms

	NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-		چله	<E>	نشستن	<E>	-	-	چله نشستن	chele neshastan	bowstring sit	sitting bowstring	living lonely and have no relation with anybody	Najafi (2000)
+	-		هو	<E>	انداختن	<E>	-	-	هو انداختن	ho andakhtan	rumor throw	throwing rumor	broadcast a wrong news	Najafi (2000)
+	-		چوب	<E>	خوردن	<E>	-	+	چوب خوردن	chob khordan	wood eat	eating wood	getting the punishment of wrong act	Najafi (2000)
-	+		چین	<E>	خوردن	<E>	-	+	چین خوردن	chin khordan	fold eat	eating fold	a small furrow or crease on a normally smooth surface	Najafi (2000)
+	-		حاجت	<E>	گرفتن	<E>	-	-	حاجت گرفتن	hajat gereftan	wish take	taking wish	wishes come true	Najafi (2000)
+	-		هول	<E>	زدن	<E>	-	-	هول زدن	hol zadan	shock hit	hitting shock	trying to get more than your needs	Najafi (2000)
+	-		حرف	<E>	داشتن	<E>	-	+	حرف داشتن	harf dashtan	talk have	having talk	objection about a particular topic	Najafi (2000)
+	-		حرکت	<E>	دادن	<E>	-	+	حرکت دادن	harkat dadan	move give	giving move	prepare for travel	Najafi (2000)
+	-		حنا	<E>	بستن	<E>	-	-	حنا بستن	nana bastan	henna close	closing henna	coloring hair with henna	Najafi (2000)
+	-		حناق	<E>	گرفتن	<E>	-	-	حناق گرفتن	honagh gereftan	diphtheria get	getting diphtheria	keep silence	Najafi (2000)
+	-		خفقان	<E>	گرفتن	<E>	-	-	خفقان گرفتن	khafeghan gereftan	suffocation get	getting suffocation	keep silence	Najafi (2000)
-	+		خاک	<E>	خوردن	<E>	+	+	خاک خوردن	khak khordan	dust eat	eating dust	be forgotten	Najafi (2000)
+	-		ختم	<E>	برداشتن	<E>	-	-	ختم برداشتن	khatm bardashtan	end pick up	picking up end	held a pray session	Najafi (2000)
+	-		خدا	<E>	داشتن	<E>	-	+	خدا داشتن	khoda dashtan	god have	having god	not being alone	Najafi (2000)
+	-		اجباری	<E>	رفتن	<E>	-	-	اجباری رفتن	ejbari raftan	compulsion go	going compulsian	going to military service	Najafi (2000)
+	-		خریدار	<E>	داشتن	<E>	-	+	خریدار داشتن	kharidar dashtan	buyer have	having buyer	somebody who attract the attention	Najafi (2000)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	خط میا	<E>	گرفتن	<E>	-	-	خستگی گرفتن	khastegi gereftan	tiredness get	getting tiredness	getting rest	Najafi (2000)
+	-	خط	<E>	انداختن	<E>	-	+	خط انداختن	khat andakhtan	line throw	throwing line	getting high	Najafi (2000)
-	+	خواب	<E>	برداشتن	<E>	-	-	خواب برداشتن	khab bardashtan	sleep pick up	picking up sleep	being highlight	Najafi (2000)
+	-	خواب	<E>	رفتن	<E>	-	-	خواب رفتن	khab raftan	sleep go	going sleep	getting numb	Najafi (2000)
+	-	گوش	<E>	خواباندن	<E>	-	-	گوش خواباندن	gosh khabandan	ear appease	appeasing ear	waiting for something	Najafi (2000)
+	-	خون	<E>	ریختن	<E>	-	+	خون ریختن	khon rikhtan	blood shed	shedding blood	doing murder	Najafi (2000)
+	-	خیال	<E>	بافتن	<E>	-	-	خیال بافتن	khial baftan	imaginati on weave	weaving imagination	having foolish and unreal wishes	Najafi (2000)
+	-	خیز	<E>	برداشتن	<E>	-	-	خیز برداشتن	khiz gereftan	jump pick up	picking up jump	ready to attack someone or something	Najafi (2000)
+	-	داغ	<E>	کردن	<E>	-	+	داغ کردن	dagh kardan	scorcher do	doing scorcher	becoming angry	Najafi (2000)
+	-	داغ	<E>	دیدن	<E>	-	-	داغ دیدن	dagh didan	scorcher see	seeing scorcher	seeing the death of family members	Najafi (2000)
+	-	دالی	<E>	کردن	<E>	-	-	دالی کردن	dali kardan	dolly do	doing dolly	showing face in sudden	Najafi (2000)
+	-	دان	<E>	پاشیدن	<E>	-	+	دان پاشیدن	dan pashidan	grain strew	strewing grain	trying to attract somebody	Najafi (2000)
+	-	دانه	<E>	ریختن	<E>	-	+	دانه ریختن	dane rikhtan	grain Shed	shedding grain	trying to attract somebody	Najafi (2000)
+	-	دبه	<E>	آوردن	<E>	-	+	دبه در آوردن	dabe daravardan	gallon evolve	evolve gallon	cheat by avoiding payment of a gambling debt	Najafi (2000)
+	-	نان	<E>	آوردن	<E>	-	+	نان در آوردن	nan daravardan	bread evolve	evolve bread	making money	Najafi (2000)
+	-	درجا	<E>	زدن	<E>	-	-	درجا زدن	darja zadan	squish hit	hitting squish	come to a stop	Najafi (2000)
+	-	چرند	<E>	گفتن	<E>	-	-	چرند گفتن	charand goftan	baloney say	saying baloney	silly idle talk	Najafi

Persian idioms

N0=:Hum	N0=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
													(2000)
+	-	دست	<E>	جنباند	<E>	-	+	دست جنباندن	dast jonbandan	hand shake	shaking hand	act or move more quickly than usual	Najafi (2000)
-	+	دست	<E>	دادن	<E>	-	+	دست دادن	dast dadan	hand give	giving hand	be prepared	Najafi (2000)
+	-	دست	<E>	داشتن	<E>	-	+	دست داشتن	dast dashtan	hand have	having hand	secret partnership in a particular activity	Najafi (2000)
+	-	دستش	<E>	آمدن	<E>	-	-	دستش آمدن	dast amadan	hand come	coming hand	become aware of something	Najafi (2000)
+	-	دست	<E>	خواندن	<E>	-	-	دست خواندن	dast khandan	hand read	reading hand	knowing about the others intentions	Najafi (2000)
+	-	دشت	<E>	کردن	<E>	-	-	دشت کردن	dasht kardan	plain do	doing plain	the advantage or benefit that is gained from selling something	Najafi (2000)
+	-	دق	<E>	آوردن	<E>	-	-	دق آوردن	degh avardan	Percussion bring	bringing percusiion	getting depression	Najafi (2000)
+	-	دل	<E>	دادن	<E>	-	-	دل دادن	del dadan	heart give	giving heart	falling in love	Najafi (2000)
+	-	گل	<E>	گفتن	<E>	-	-	گل گفتن	gol goftan	flower say	saying flower	telling the truth	Najafi (2000)
+	-	دل	<E>	داشتن	<E>	-	+	دل داشتن	del dasht	heart have	having heart	being brave	Najafi (2000)
+	-	دل	<E>	کندن	<E>	-	-	دل کندن	del kandan	heart dig	dig heart	leave and never return to something	Najafi (2000)
+	-	دم	<E>	گرفتن	<E>	-	-	دم گرفتن	dom gereftan	tail get	getting tail	standing in a line behind each other	Najafi (2000)
+	-	دماغ	<E>	گرفتن	<E>	-	-	دماغ گرفتن	damagh gereftan	nose get	grtting nose	becoming upset because of unimportant thing	Najafi (2000)
+	-	دود	<E>	گرفتن	<E>	-	+	دود گرفتن	dod gereftan	smoke get	getting Smoke	smoking opium	Najafi (2000)

Persian idioms

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	دوره	<E>	افتادن	<E>	-	-	دوره افتادن	dore oftadan	period fall	falling period	for doing a particular thing going to different places and asking different people	Najafi (2000)
+	-	دوره	<E>	داشتن	<E>	-	-	دوره داشتن	dore dashtan	periodicity have	having periodicity	attending in a successive parties	Najafi (2000)
+	-	ذات	<E>	داشتن	<E>	-	+	ذات داشتن	zat dashtan	essence have	having essence	being good human being	Najafi (2000)
+	-	فضا	<E>	رفتن	<E>	-	+	فضا رفتن	faza raftan	space go	going space	being high	Najafi (2000)
+	-	کپ	<E>	کردن	<E>	-	-	کپ کردن	kop kardan	demijohn do	doing demijohn	being shocked	Najafi (2000)
+	-	سرویس	<E>	کردن	<E>	-	+	سرویس کردن	servis kardan	service do	doing service	annoying someone	Najafi (2000)
+	-	آفت	<E>	خوردن	<E>	-	-	آفت خوردن	afat khordan	pest eat	eating pest	when the police suspect and stop you	Najafi (2000)
+	-	آف	<E>	کردن	<E>	-	+	آف کردن	af kardan	off do	doing off	being quiet	Najafi (2000)
+	-	گه	<E>	خوردن	<E>	-	+	گه خوردن	goh khordan	shit eat	eating shit	saying bullshit	Najafi (2000)
+	-	کف	<E>	کردن	<E>	-	+	کف کردن	kaf kardan	foam do	doing foam	a feeling of amazement and admiration, caused by something beautiful, remarkable, or unfamiliar	Najafi (2000)
+	-	بخت	<E>	برگشتن	<E>	-	-	بخت برگشتن	bakht bargashtan	fortune return	returning fortune	subjected to or marked by misfortune	Najafi (2000)
+	-	قفل	<E>	کردن	<E>	-	+	قفل کردن	ghofl kardan	lock do	doing lock	being shocked	Najafi (2000)
+	-	خایه	<E>	کردن	<E>	-	-	خایه کردن	khaye kardan	testicle paste	pasting testicle	being afraid	Najafi (2000)

NO=:Hum	NO=:N-hum	C1	Det	Verb	Modify	Passive	Literal	Persian idioms	Transliteration	Gloss	Rough translation	Exact meaning	Reference
+	-	خایه	<E>	چسباندن	<E>	-	-	خایه چسباندن	khaye chasbandan	testicle do	doing testicle	being afraid	Najafi (2000)
+	-	کس	<E>	گفتن	<E>	-	-	کس گفتن	kos goftan	pussy say	saying pussy	saying bullshit	Najafi (2000)

## Appendix B. Dictionary of the inflected forms (*kard*, *kon* “to do”)

V:HP2s. کرده ای, کرد	V:HF3s. می کرده, کرد	Neg+V:GS1s. نکرده, کرد	Neg+V:HF3s. است, کرد
V:HP2p. کرده اید, کرد	V:HC1s. می کن, کن	Neg+V:GS2p. نکرده, کرد	Neg+V:HF1s. نمی کرده ام, کرد
V:HP1p. کرده ایم, کرد	V:HC2s. می کنی, کن	Neg+V:GS2s. نکرده, کرد	Neg+V:HF3p. نمی کرده اند, کرد
V:HP3s. کرده است, کرد	V:HC2p. می کنید, کن	Neg+V:GS3p. نکرده, کرد	Neg+V:HF3s. نمی کرده, کرد
V:HP1s. کرده ام, کرد	V:HC1p. می کنیم, کن	Neg+V:GS3s. نکرده, کرد	Neg+V:HC2s. نمی کنی, کن
V:HP3p. کرده اند, کرد	V:HC3s. می کند, کن	Neg+V:HP2s. نکرده ای, کرد	Neg+V:HC2p. نمی کنید, کن
V:HP3s. کرده, کرد	V:HC3p. می کنند, کن	Neg+V:HP2p. نکرده اید, کرد	Neg+V:HC1p. نمی کنیم, کن
V:Hr. بکن, کن	V:GC3s. میگرد, کرد	Neg+V:HP1p. نکرده ایم, کرد	Neg+V:HC3s. نمی کند, کن
V:HU2s. بکنی, کن	V:GC2s. میگردی, کرد	Neg+V:HP3s. نکرده است, کرد	Neg+V:HC1s. نمی کنم, کن
V:Hr. بکنی, کن	V:GC2p. میگردید, کرد	Neg+V:HP1s. نکرده ام, کرد	Neg+V:HC3p. نمی کنند, کن
V:HU2p. بکنید, کن	V:GC1p. میگردیم, کرد	Neg+V:HP3p. نکرده اند, کرد	Neg+V:GC3s. نمیگرد, کرد
V:HU1p. بکنیم, کن	V:GC1s. میگردم, کرد	Neg+V:HP3s. نکرده, کرد	Neg+V:GC2s. نمیگردی, کرد
V:HU3s. بکنند, کن	V:GC3p. میگردند, کرد	Neg+V:HS2s. نکنی, کن	Neg+V:GC2p. نمیگردید, کرد
V:HU1s. بکنم, کن	V:HF2s. میگرده ای, کرد	Neg+V:HS2p. نکنید, کن	Neg+V:GC1p. نمیگردیم, کرد
V:HU3p. بکنند, کن	V:HF2p. میگرده اید, کرد	Neg+V:HS1p. نکنیم, کن	Neg+V:GC1s. نمیگردم, کرد
V:GC3s. می کرد, کرد	V:HF1p. میگرده ایم, کرد	Neg+V:HS3s. نکنند, کن	Neg+V:GC3p. نمیگردند, کرد
V:GC2s. می کردی, کرد	V:HF3s. میگرده است, کرد	Neg+V:HS1s. نکنم, کن	Neg+V:HF2s. نمیگرده ای, کرد
V:GC2p. می کردید, کرد	V:HF1s. میگرده ام, کرد	Neg+V:HS3p. نکنند, کن	Neg+V:HF2p. نمیگرده اید, کرد
V:GC1p. می کردیم, کرد	V:HF3p. میگرده اند, کرد	Neg+V:GC3s. نمی کرد, کرد	Neg+V:HF1p. نمیگرده ایم, کرد
			نمیگرده
V:GC1s. می کردم, کرد	V:HF3s. میگرده, کرد	Neg+V:GC2s. نمی کردی, کرد	Neg+V:HF3s. است, کرد
V:GC3p. می کردند, کرد	V:HC2s. میکنی, کن	Neg+V:GC2p. نمی کردید, کرد	Neg+V:HF1s. نمیگرده ام, کرد
V:HF2s. می کرده ای, کرد	V:HC2p. میکنید, کن	Neg+V:GC1p. نمی کردیم, کرد	Neg+V:HF3p. نمیگرده اند, کرد
V:HF2p. می کرده اید, کرد	V:HC1p. میکنیم, کن	Neg+V:GC1s. نمی کردم, کرد	Neg+V:HF3s. نمیگرده, کرد
V:HF1p. می کرده ایم, کرد	V:HC3s. می کند, کن	Neg+V:GC3p. نمی کردند, کرد	Neg+V:HC2s. نمیکنی, کن
می کرده			
V:HF3s. است, کرد	V:HC1s. میکنم, کن	Neg+V:HF2s. نمی کرده ای, کرد	Neg+V:HC2p. نمیکنید, کن
V:HF1s. می کرده ام, کرد	V:HC3p. می کنند, کن	Neg+V:HF2p. نمی کرده اید, کرد	Neg+V:HC1p. نمیکنیم, کن
V:HF3p. می کرده اند, کرد	Neg+V:GS1p. نکرده, کرد	Neg+V:HF1p. نمی کرده ایم, کرد	Neg+V:HC3s. نمیکنند, کن
			Neg+V:HC1s. نمیکنم, کن
			Neg+V:HC3p. نمیکنند, کن

## Appendix C. Dictionary of the inflected forms (*زد, zad, زن, zan* “to hit”)

V:Hr.زن, بزن	V:HF3s.می زده, زد	Neg+V:GS1s.نزد, نزد	نمی زده است, زد, HF3s. Neg+V:
V:HU2s.بزنی, زن	V:HC2s.می زنی, زن	Neg+V:GS2p.نزد, نزد	نمی زده ام, زد, HF1s. Neg+V:
V:Hr.بزنی, زن	V:HC2p.می زنی, زد	Neg+V:GS2s.نزد, نزد	نمی زده اند, زد, HF3p. Neg+V:
V:HU2p.بزنید, زن	V:HC1p.می زنی, زد	Neg+V:GS3p.نزد, نزد	نمی زده, زد, HF3s. Neg+V:
V:HU1p.بزنیم, زن	V:HC3s.می زنی, زد	Neg+V:GS3s.نزد, نزد	نمی زنی, زن, HC2s. Neg+V:
V:HU3s.بزند, زن	V:HC1s.می زنی, زد	Neg+V:HP2s.نزده ای, زد	نمی زنی, زد, HC2p. Neg+V:
V:HU1s.بزنم, زن	V:HC3p.می زنی, زد	Neg+V:HP2p.نزده ای, زد	نمی زنی, زد, HC1p. Neg+V:
V:HU3p.بزنند, زن	V:GC3s.میزد, زد	Neg+V:HP1p.نزده ایم, زد	نمی زنی, زد, HC3s. Neg+V:
V:HP2s.زده ای, زد	V:GC2s.میزدی, زد	Neg+V:HP3s.نزده است, زد	نمی زنی, زد, HC1s. Neg+V:
V:HP2p.زده ای, زد	V:GC2p.میزدید, زد	Neg+V:HP1s.نزده ام, زد	نمی زنی, زد, HC3p. Neg+V:
V:HP1p.زده ایم, زد	V:GC1p.میزدیم, زد	Neg+V:HP3p.نزده اند, زد	نمی زد, زد, GC3s. Neg+V:
V:HP3s.زده است, زد	V:GC1s.میزدم, زد	Neg+V:HP3s.نزده اند, زد	نمی زد, زد, GC2s. Neg+V:
V:HP1s.زده ام, زد	V:GC3p.میزدند, زد	Neg+V:HS2s.نزنی, زن	نمی زدید, زد, GC2p. Neg+V:
V:HP3p.زده اند, زد	V:HF2s.میزده ای, زد	Neg+V:HS2p.نزنی, زد	نمی زدیم, زد, GC1p. Neg+V:
V:HP3s.زده, زد	V:HF2p.میزده ای, زد	Neg+V:HS1p.نزنیم, زن	نمی زد, زد, GC1s. Neg+V:
V:GC3s.می زد, زد	V:HF1p.میزده ایم, زد	Neg+V:HS3s.نزند, زن	نمی زدند, زد, GC3p. Neg+V:
V:GC2s.می زدی, زد	V:HF3s.میزده است, زد	Neg+V:HS1s.نزنم, زن	نمی زده ای, زد, HF2s. Neg+V:
V:GC2p.می زدید, زد	V:HF1s.میزده ام, زد	Neg+V:HS3p.نزنند, زن	نمی زده ای, زد, HF2p. Neg+V:
V:GC1p.می زدیم, زد	V:HF3p.میزده اند, زد	Neg+V:GC3s.نمی زد, زد	نمی زده ایم, زد, HF1p. Neg+V:
V:GC1s.می زد, زد	V:HF3s.میزده, زد	Neg+V:GC2s.نمی زدی, زد	نمی زده است, زد, HF3s. Neg+V:
V:GC3p.می زدند, زد	V:HC2s.میزنی, زن	Neg+V:GC2p.نمی زدید, زد	نمی زده ام, زد, HF1s. Neg+V:
V:HF2s.می زده ای, زد	V:HC2p.میزنی, زد	Neg+V:GC1p.نمی زدیم, زد	نمی زده اند, زد, HF3p. Neg+V:
V:HF2p.می زده ای, زد	V:HC1p.میزنیم, زن	Neg+V:GC1s.نمی زد, زد	نمی زده, زد, HF3s. Neg+V:
V:HF1p.می زده ایم, زد	V:HC3s.میزند, زن	Neg+V:GC3p.نمی زدند, زد	نمی زنی, زن, HC2s. Neg+V:
V:HF3s.می زده است, زد	V:HC1s.میزنم, زن	Neg+V:HF2s.نمی زده ای, زد	نمی زنی, زد, HC2p. Neg+V:
V:HF1s.می زده ام, زد	V:HC3p.میزند, زن	Neg+V:HF2p.نمی زده ای, زد	نمی زنی, زد, HC1p. Neg+V:
V:HF3p.می زده اند, زد	Neg+V:GS1p.نزد, نزد	Neg+V:HF1p.نمی زده ایم, زد	نمی زنی, زد, HC3s. Neg+V:
			نمی زده, زد, HC1s. Neg+V:
			نمی زنی, زد, HC3p. Neg+V: