Aalto University
School of Science
Master's Programme in Machine Learning and Data Mining

Murtaza Hazara

Unsupervised methods in multilingual and multimodal semantic modeling

Master's Thesis

Espoo, August 4, 2014

Supervisor: Prof. Erkki Oja, Aalto University

Instructor: Prof. Timo Honkela, Aalto University

Aalto University School of Science Master's Programme in Machine Learning and Data Mining		ABSTRACT OF THE MASTER'S THESIS
Author: Murtaza Hazara		
Title: Unsupervised methods in multilingual and multimodal semantic modeling		
Number of pages: 101	Date: September 15, 2014	Language: English
Professorship: Information and Computer Science		Code: T-61
Supervisor: Prof. Erkki Oja		
Advisor: Prof. Timo Honkela		

Abstract:

In the first part of this project, independent component analysis has been applied to extract word clusters from two Farsi corpora. Both word-document and word-context matrices have been considered to extract such clusters. The application of ICA on the word-document matrices extracted from these two corpora led to the detection of syntagmatic word clusters, while the utilization of word-context matrix resulted in the extraction of both syntagmatic and paradigmatic word clusters. Furthermore, we have discussed some potential benefits of this automatically extracted thesaurus.

In such a thesaurus, a word is defined by some other words without being connected to the outer physical objects. In order to fill such a gap, symbol grounding has been proposed by philosophers as a mechanism which might connect words to their physical referents. From their point of view, if words are properly connected to their referents, their meaning might be realized. Once this objective is achieved, a new promising horizon would open in the realm of artificial intelligence.

In the second part of the project, we have offered a simple but novel method for grounding words based on the features coming from the visual modality. Firstly, indexical grounding is implemented. In this naïve symbol grounding method, a word is characterized using video indexes as its context. Secondly, such indexical word vectors have been normalized according to the features calculated for motion videos. This multimodal fusion has been referred to as the pattern grounding. In addition, the indexical word vectors have been normalized using some randomly generated data instead of the original motion features. This third case was called randomized grounding. These three cases of symbol grounding have been compared in terms of the performance of translation. Besides that, word clusters have been excerpted by comparing the vector distances and from the dendrograms generated using an agglomerative hierarchical clustering method.

We have observed that pattern grounding exceled the indexical grounding in the translation of the motion annotated words, while randomized grounding has deteriorated the translation significantly. Moreover, pattern grounding culminated in the formation of clusters in which a word fit semantically to the other members, while using the indexical grounding, some of the closely related words dispersed into arbitrary clusters.

Keywords: symbol grounding, automatic thesaurus extraction, multimodal fusion, hierarchical clustering

Acknowledgment

First and foremost, I would like to express my sincere gratitude to my instructor Prof. Timo Honkela for giving me an opportunity to work on this lovely evergreen subject and supporting me by valuable instructions and his kind behavior. In addition, I would like to thank my supervisor Prof. Erkki Oja for his motivation, enthusiasm, and immense knowledge. Besides my supervisor and instructor, my special thanks goes to Klaus Förger whose guidance and help initiated my motivation. Finally, I appreciate all those people who participated in annotating the motion videos; without their help, this thesis would not have been possible.

Table of contents

1. Introduction

With the reduction in Internet connection prices, the utilization of more powerful computer memory and faster broadband services in every corner of the globe, a vast collection of text documents have been rapidly amassed. This has urged industrial and governmental corporations to concentrate on creating intelligent agents for extracting valuable information and knowledge from text corpora. Such intelligent agents can be helpful in various decision making processes such as monitoring social networks for marketing or anticipating potential terrorist attacks. In the realm of artificial intelligence, an apt classical question is how we can enable such artificial intelligence agents to understand natural languages.

Identification of similar words has been one of the traditional but rudimentary approaches toward automatic semantic analysis [127]. There are many ways of expressing one concept, and many concepts can be expressed by one word. In the last few decades, several algorithms have been proposed for an automatic construction of thesaurus which can capture such many-to-many semantic relationships among words. One can benefit from an automatic thesaurus for information retrieval tasks and overcoming the data sparsity problem. It can also shed light on the content of a large corpus. However, from the semantic point of view, in a thesaurus, a word is just defined by or connected to other words without being linked to the outer objects. On the other hand, symbol grounding is a more sophisticated approach for conquering the meaning of a word.

Symbol grounding can be interpreted as a hook that connects the words in our heads to the outer objects. In anthropic mechanism [44], it is believed that cognition can be explained by physical rules. Furthermore, Fodor [125] suggested that the meaning of a symbol is grounded in the relationship between the symbol system and the world. If correct, one can stimulate cognition by providing an artificial intelligence with the right rules for manipulating symbols. In other words, we would be able to create an artificial agent that can understand natural languages such as English as we do. Although symbol grounding and consciousness are quite popular concepts among philosophers, they have not gained considerable recognition by information scientists. In this project, we study the influence of symbol grounding on automatic translation and synonym detection, which can be regarded as one of the first works in the translation field according to our best knowledge.

In the first part of this project, WordICA is applied on two Farsi corpora. This part is mainly about extracting a numerical representation for every word in Farsi corpora. Numerical representations of the words can be obtained by defining words in the vector space model. In this model, a word is demonstrated by a vector whose components are real numbers. For example, a word can be mapped to an integer number by counting its frequency in a specific document, and if we consider all the available documents, then we can form a vector in a d dimensional space, where d denotes the number of documents. This kind of numerical representation can then be input to independent component analysis (ICA) algorithm for extracting automatic thesaurus. Furthermore, Independent component analysis has been applied for word clustering because it has the capability of automatic document clustering which can be utilized in future works.

In the second part of the project, we proposed a novel framework for capturing the meaning of motion-related words. Arena made by OptiTrack was employed to create motion videos. In addition, motion features such as the means and standard deviations of coordinates, velocities, and the accelerations of different body parts were calculated. Every motion video was represented by such 602-dimensional motion data.

An online web-based form has been designed where English, Finnish, Swedish, and Farsi speaking people can annotate a set of motion videos. Word vectors have been formed by counting the frequency of a word in a video. There are 124 motion videos; hence, words are represented by 124 dimensional vectors. Then, the word vectors were normalized by the 602-dimensional motion data. As a matter of fact, we want to study how the system performance would change in terms of translation and synonym detection, if we fuse textual and visual modalities. In other words, words are grounded by the data from visual modality. That's how this part of the project is related to the symbol grounding problem.

Since Farsi is one the studied language in this project, and some readers might not be so familiar with it, a very brief introduction to this language is presented in section 1.1. Section 1.2 explains the symbol grounding problem very shortly. Section 1.3 expounds multimodal technologies. Section 1.5 mentions some application of text mining. Then, in Chapter 2, the methodologies including PCA, ICA, WordICA, and hierarchical clustering is stated. Besides that, in this chapter, the methods of synonyms detection and translation of motion-related words are clarified. Chapter 3 describes the results of both the symbol grounding and word clustering experiments. Finally, Chapter 4 asserts a conclusive statement about the exploitation of WordICA for word clustering and symbol grounding for automatic translation systems.

1.1. Persian (Farsi) language

Persian is an Iranian language which is itself a branch of Indo-European languages [131, 132, 133, 135]. It is a polycentric language spoken by almost 130 million people mainly in Iran, Afghanistan, and Tajikistan. It is also spoken in Iraq, Pakistan, Uzbekistan, and Turkmenistan by minorities. It is natively referred to Farsi in Iran. Although Persian is the official language in Iran, nearly half of its population is non-native Persian speaker; the other languages spoken in Iran are Azerbaijani, Balochi, Kurdish, and Arabic [134]. In Afghanistan, it is known as Dari, an official language together with Pashtu [136].

Old Persian was written from left to right on cuneiform script. After the Islamic conquest of Iran, in the ninth century, modern Persian was established and enriched by many Arabic loanwords. Its alphabet is based on Arabic script with four more letters being added to it. Accordingly, modern Persian is written from right to left. Furthermore, there is no difference between capital and lower case letters.

It is a morphologically rich language, and there are over a hundred affixes to form new words. It is claimed that Farsi is an agglutinative language. New words can be formed both by combining bounding morphemes and compounding two existing words. In writing, Farsi generally makes use of only consonants and long vowels, not showing short vowels in the written form. In addition, Professor Mahmoud Hessaby demonstrated that Persian can derive more than 226 million words [137].

1.1.1. Persian adjectives and adverbs

Persian adjectives have a limited inflection space; they may be simple, comparative, or superlative. If a simple adjective is not an Arabic loanword, its comparative and superlative form can be easily made by adding a suffix to it.

Table 1: Persian comparative and superlative adjectives

Simple adjective	Comparative adjective	Superlative adjective
(bad) بد-bad	(worse)بدتر-badtar	(the worst)بدترین-badtarin
(wild)خشن-khashen	(wilder)خشن‌تر-khashentar	(the wildest)خشن‌ترین-khashentarin
(kind)مهربان-mehraban	(kinder)مهربان‌تر-mehrabantar	(the kindest)مهربان‌ترین-mehrabantarin

Adverbs are primarily identical with, or derived from, nouns or adjectives. Some Arabic loan words are transformed into a manner or a sentence adverb by adding the Arabic "tanwin accusative" loans ending in –an: e.g., ghalban-قلباً (meaning by heart). Adjectival adverbs, which are identical with or originating from adjectives, are mostly of manner: تند-tond (quickly, hastily). Intensifying adverbs, which qualify other adjectives or adverbs, are the quantifying adjectives: خیلی-kheyli (very).

Table 2: Persian Adverbs

adverb	meaning	Type of adverb	root	Part of speech
(nesbatan)نسبتاً	Relatively	Sentence adverb	(relation)نسبت-nesbat	Noun
(nazdik)نزدیک	Near	Nominal adverb	(near, now)نزد-nazd	Noun
(yawash)یواش	Slowly	Adjectival adverb	(slow, slowly)یواش-yawash	Adjective, adverb
(besiyar)بسیار	very	Intensifying adverb	(very)بسیار-besiyar	Adjective

1.1.2. Zero-width-non-joiner problem

Detection of word boundaries is an important preprocessing task in statistical text mining. In the literature, this task is referred to as tokenization. Tokenization of Persian text documents is still a challenging process as some words might be written in a concatenated form, or attached morpheme might be separated with a space.

In Farsi, the shape of a letter is determined by whether it is joined or separate. For example, the verb "is doing" in Farsi might be written in these three different forms:

1. می‌کند
2. می کند
3. میکند

However, only the first version is grammatically correct. There are two morphemes in this verb:

1. می (mi)= it is a morpheme that shows present continuous terms in this case
2. کند (konad)= it is the root of verb "do" for the present tense

When there is space between two things, it means that they are two words. Thus, two morphemes that are parts of the same word should not be separated by a space. On the other hand, if they are joined, its spelling would not be correct according to the Farsi grammar. The solution is to use a zero-width-non-joiner; in this case, while there is no space between the last letter of the first morpheme and the first letter of the second morpheme, they are not joined.

In the first part of this project, frequencies of words should be extracted to convert words into a numerical representation. When counting the frequencies of substrings in Farsi text, the zero-width-non-joiner (ZWNJ) can be problematic. Zero-width-non-joiner is a kind of space that cannot be seen, and it is there to separate letters before and after without adding a space. It is most frequent in verbs where its derivational morpheme should be separated without a space. In addition, typing Farsi words with English letters is quite popular, which has worried Persian linguist experts. Thus, there was an extensive effort to utilize machine learning techniques for building online user friendly websites in which a user can type Farsi words using Latin letters, and receive the words back in Perso-Arabic letters. Behnevis [42] is an online service which offers such conversion. The problem of such online tools is that too ZWNJ might be added without the user being aware of it. For example, **both of these two verbs seem to be correct and equivalent of each other**:

1. می‌کند=(U+0645 U+06CC U+200C U+06A9 U+0646 U+062F)
2. می‌کند=(U+0645 U+06CC U+200C **U+200C U+200C** U+06A9 U+0646 U+062F)

However, it is only by comparing their Unicode that we can notice their differences. In fact, in the first version, only one ZWNJ (U+200C) is used, while in the second version, two more ZWNJs are added. Hence, when working with Farsi words, unnecessary ZWNJs must be removed in the preprocessing steps; otherwise, frequencies of some words might not be captured correctly, which results in inaccurate numerical representations of words.

1.2. Symbol grounding problem

A symbol is any object which is a part of a symbol system such as natural language [50-54]. English alphabet is an example of a set of symbols whose shapes are selected arbitrarily. In addition, the shape of a symbol is neither the hint of its meaning nor the shape of the object it refers to. A tool is required to find the referent of a symbol, and when the symbols are connected to their referent, they become meaningful. Symbol grounding is a mechanism using which the object to which a symbol refers can be detected.

Symbol grounding problem is about how words get their meaning. This problem can be further clarified using an intuitive example. Suppose that you want to learn a foreign language such as Chinese, and the only available tool is a monolingual Chinese dictionary in which the Chinese words are defined using other Chinese words; you cannot find any kind of clue such as an image in this dictionary. In other words, the words in this dictionary are not connected to anything in the world. In this case, when you intend to recognize the meaning of a new word, you have to also look up all the words which are included in the definition of that word; this will make your search for the meaning of a word like an infinite loop. The only reason cryptologists of ancient languages and secret codes seem to be able to successfully accomplish something very like this is that their efforts are grounded in a first language and in real world experience and knowledge [54].

The aforementioned example has been inspired by the famous Chinese room argument proposed by John Searle [143]. He has imagined himself alone in a room communicating appropriately with people outside the room in Chinese just by following the computer instructions which manipulates the Chinese symbols, while in fact, he has no knowledge of Chinese. Although people who are outside think that he is a Chinese speaker, he did not manage to realize the meaning of any of the Chinese words. He concludes that understanding is a biological process and no computer can understand Chinese by following a program. In other words one cannot get semantics from syntax. Searle's argument has initiated a hot debate among recent philosophers. Critics have offered several counter arguments among which "The Robot Reply" captured our attention [144]. In this reply, it is conceded that a natural language processing program does not create any understanding, but if the program is embedded inside a robotic body with sensors and motor enabling the robot to make contact with the physical entities, it would be possible for the robot to understand a natural language. That's how we are inspired to utilize multimodal language technologies to capture the semantics of motion related words.

Symbolists like Fodor believe that symbols get their meaning by being appropriately connected to the outer objects in the World [48, 49]. The fact that our own symbols do have intrinsic meaning whereas the computer's symbols do not indicates that automatic translation and artificial intelligence in general can benefit from symbol grounding.

The symbol grounding problem is also relevant to consciousness. According to Max Velmans and Susan Schneider [43], consciousness is the most familiar and the most mysterious aspect of our lives. There are two doctrines concerning consciousness. In anthropic mechanism, everything about human beings can be explained in mechanical terms as surely as can everything about clockwork or gasoline engines [44]. However, one of the chief obstacles that all mechanistic theories have faced is providing a mechanistic explanation of the human mind and consciousness. For one, although Descartes endorsed profoundly the mechanistic conception of the material world and some human function such as passion, he argued that one cannot explain the conscious mind in terms of the spatial dynamics. The theory that opposed to this mechanism is vitalism, which maintains that vital activities cannot be explained by the laws which govern lifeless matter; hence, consciousness can be neither explained nor measured by physical laws. It is not of our major concern whether consciousness can ever be explained mechanistically. However, the question of how consciousness might be related to language is more pertinent to this study.

Ned Block proposes two distinct types of consciousness which he called phenomenal (P-consciousness) and access (A-consciousness) [45]. P-conscious states include the experiential states we have when we see, hear and have pains. These experiences, considered independently of any impact on behavior. A-consciousness, on the other hand, is the phenomenon whereby information in our minds is accessible for verbal report, reasoning, and the control of behavior. So, when we perceive, information about what we perceive is access conscious; when we introspect, information about our thoughts is access conscious [46], and so on. According to David Chalmers, A-consciousness can be understood in mechanistic terms [47]. Accordingly, it would be possible to unravel the mystery of relation between perceiving an event and its verbal report.

1.3. Multimodal technology

In a multimodal technology the communication is accomplished through various modalities. A mode or modality refers to a channel through which a message or information is sent or received. For examples, in human-human communication, multifarious modalities such as speaking, writing, gesturing and touching are exploited. Similarly, multimodality can be seen in human-computer interaction by the use of different input/output channels. Exploiting multiple modalities enhances our human-computer interaction and makes it more natural.

Nowadays, multimodal systems can be found in sophisticated inventions such as haptic devices for generating sensation to the skin [56, 57], disability assistive applications [58], wearable virtual devices [59, 60], GPS navigation systems [61], military super-soldier enhancing devices [62], personal digital assistants [63], speech enabled interfaces [64], smart advertisement [65, 66], virtual keyboard [67], virtual reality [68, 69], and many other technologies [55].

The central process in a multimodal system is multimodal fusion which refers to assimilating information from assorted input channels. In addition to improving interface design, fusion of multiple modalities can also increase the accuracy of classification and decision making processes. For example, in [111], the authors have combined facial and voice data to recognize four emotions: "sadness, anger, happiness, and neutral state". Their results demonstrate that facial data is more informative about emotion than acoustic data. In fact, according to their experiments, the performance of their emotion recognition system was 70.9 percent when the acoustic data was the only utilized input, while exploiting facial data led to an overall performance of 85 percent. On the other hand, by fusion of voice and facial data, the overall performance of the recognition system soared to 89.1 percent.

The abundant successful stories of multimodal fusion have propelled research scientists to focus more on this subject. For example, in [70], the fusion of audio-visual features along with other textual information has been shown to be effective in detecting events from a team sport video.

There are three approaches for combining different modalities: early fusion, late fusion, and the hybrid method. In early fusion, the information is fused at the feature level; for instance, audio and visual feature vectors are concatenated for a classification task. In the literature, this approach is known as the recognition-based fusion [71-73]. In late fusion, multiple modalities are connected in the semantic space. For instance, audio and visual feature vectors can be processed by some classifiers to produce higher level representations such as phoneme and viseme; then, instead of combining audio and visual feature vectors, phoneme and viseme vectors can be incorporated. Late fusion is also referred to as decision level fusion [74-80]. Hybrid method is a combination of early and late fusion [81-85]. In [66], an overview of these three methods and an in-depth view of various strategies for multimodal fusion is offered.

1.3.1. Multimodal language technology

Multimodal language technologies refer to technologies that combine text or audio features with features from some other modalities such as image. The majority of researches in this field have focused on combining audio and vision for the various purposes including but not limited to speech recognition [87-90], biometric identification [91, 92], speaker recognition [93, 94], semantic concept detection [90, 93], video classification [96], and human tracking [97].

Some researchers have also tried to merge textual information with features from other modalities. In [98], the authors have applied maximum entropy model to fuse text with image based data at the feature level for semantic image indexing. In [99], features from audio, video, text, and weblog modalities have been fused at the hybrid level for the purpose of sport video analysis. In [100], features from text, audio, video and speech modalities have been combined at the hybrid level for video topic clustering. In [101], features from text (closed caption), audio, video and motion modalities have been merged linearly at the decision level for video retrieval.

1.3.2. Automatic multimodal translation

The majority of the previous works in the realm of multimodal language technology has been concentrated on information retrieval, clustering and classification. There exist few researches which have been devoted to automatic multimodal translation. For example, in [102], the authors have been trying to develop a system so that the lip movements in dubbed movies can be automatically synchronized to the translated speech. In [103], Duygulu et al have built an automatic lexicon from a set of annotated images; these images were annotated in different languages including English and French. Each image annotation consists of a set of words. Every word can be linked to a region in an image. There is no information about such connection between the words in an annotation and the regions in the corresponding image. This connection was learned by applying a variant of EM algorithm. Then each region in an image is linked to words from various languages. Accordingly, by object recognition, an automatic lexicon was created. In [104], textual and visual features were fused for automatic annotation, and from the authors' point of view, the problem of image annotation could be viewed as analogous to the problem of cross-lingual retrieval.

1.4. Motion capture

Motion capture is the process of recording a live motion and translating it into actionable data that allows a 3D recreation of the performance. It involves measuring an object's position and orientation in physical space, then recording that information in a computer-usable form. Objects of interest include human and non-human bodies, facial expressions, camera or light positions, and other elements in a scene [138]. Data acquisition is implemented using markers attached near the joints of an actor; using these markers, low level data such as the positions and angles between the markers are recorded. Such raw low level data enables us to animate a humanoid character. Motion capture was started as a photogrammetric tool analysis in 1970's; later, it found its way into education, training, sports, biomechanics, and gesture recognition, and nowadays, it is extensively exploited in generating animation for cinema and video games [139, 140].

1.5. Text mining

Text mining is a data mining technique for extracting useful information from unstructured or semi-structured text documents. Text mining is an interdisciplinary field which incorporates data mining, web mining, information retrieval, information extraction, computational linguistics and natural language processing. Some of the basic tasks in text mining include document classification, document clustering, concept co-occurrence, named entity recognition, part-of-speech tagging and summarization. Text mining has been exploited in multifarious fields such as security [105], biomedical [106-109], marketing [110, 111], sentiment analysis [112, 113], business intelligence [114], and social media monitoring [115, 116].

1.5.1. Word clustering

Word clustering refers to the task of automatically identifying semantically similar words and putting them in the same group. One of its immediate applications is the automatic construction of thesaurus. Automatically produced thesaurus is a requisite part of translation and information retrieval systems. According to Dekang Lin [117], a certain sense of a word might have been used at a specific period of time, which makes it unlikely to be captured by manually compiled lexicons. For example, by analyzing San Jose Mercury corpus (45 million words), it has been found that more than half of the occurrences of the word “westerner” refers to “hostage”. Thus, when searching for hostage-related articles, westerner might be a good keyword search candidate. Overcoming the data sparsity problem is another advantage of automatic word clustering; Dagan has demonstrated in [118], that smoothing the maximum likelihood estimate of a word based on the likelihood of its synonym has exceled the back-off smoothing. In [122, 123], it is stated that word clusters can shed light on the overall content of underlying corpus; in other words, word clusters can be viewed as a summary of a large corpus. In [124] a broad overview of clustering algorithms has been given. Furthermore, Implementation of some text clustering algorithm can be found in several toolkits such as Lemur [120] and BOW toolkit in [121].

1.5.2. Syntagmatic vs. paradigmatic

An automatic thesaurus might capture various semantic relations such as synonym, antonym, hyponym, meronym and hypernym. Sahlgren [127] has placed various semantic relations among words under two umbrellas namely syntagmatic and paradigmatic. Words co-occuring in the text are syntagmatically related. Such relation is linear and combinatorial; in other words, words with this kind of relation can be combined together. For instance, "shoot" and "gun" are syntagmatically related. Another example of such relation is the words which emerge in a normal sentence like "I am hungry." On the other hand, in a paradigmatic relation, words can be substituted. Such relations hold between words that do not co-occur in the same context but whose neighboring words are often the same, like the words "house" and "apartment" in the sentence "I own [an apartment | a house]". These two relations are often regarded as orthogonal axes in a grid.

	Paradigmatic relations Selections: "x or y or..."			
Syntagmatic relations	She	buys	green	paint
Combinations:	He	draws	blue	clay
"x and y and..."	They	paint	red	color

2. Methodologies

Principal component analysis is applied on the high dimensional dataset collected from motion videos to convert them to two dimensional data points so that they can be visualized. In addition, WordICA which is based on Independent component analysis is employed on word-word and word-document matrix for extracting word features which can be utilized for automatic detection of word clusters. In the following sections, I have explained these methods briefly.

2.1. Principle component analysis

Principal component analysis is an eigenvector-based procedure which converts a set of correlated observations to a set of uncorrelated variables. PCA was first introduced by Karl Pearson [38]. It is a popular statistical tool utilized for exploratory data analysis and making predictive models. This tool can reveal the underlying structure of data in a way that best explains the variance in data. When encountering a high dimensional data, a preprocessing step is to visualize its reduced dimensional version. PCA can be very useful by yielding such a lower dimensional picture of the high dimensional data. Such dimensionality reduction can be achieved by projecting the data into the first few principal components. In addition to dimensionality reduction, it can be applied to reduce the noise and prevent the overfitting problem. PCA is closely related to the factor analysis and singular value decomposition.

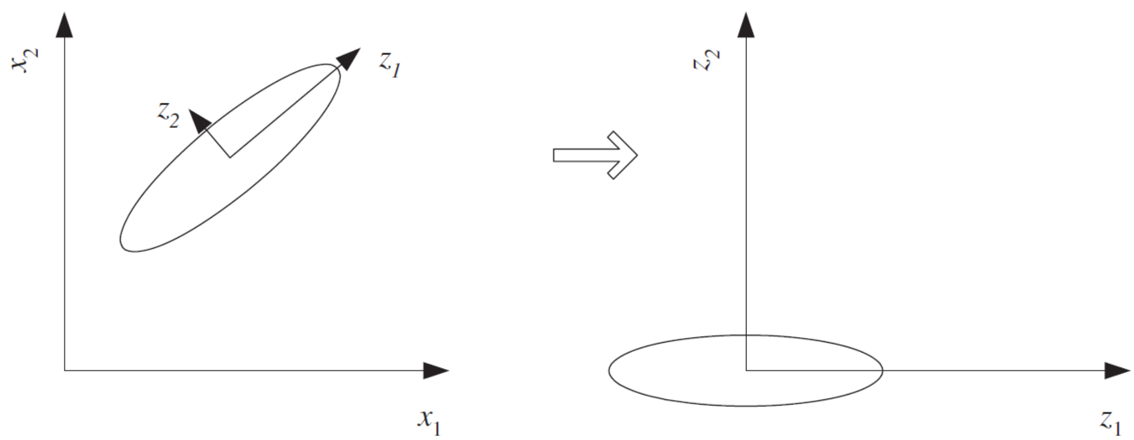


Figure 1: Principal components analysis centers the sample and then rotates the axes to line up with the directions of highest variance. If the variance on $\boldsymbol{z_2}$ is too small, it can be ignored and we have dimensionality reduction from two to one.

Figure 1 depicts a very simple example of PCA when applied on a two dimensional dataset. The principal component, z_1, is along the direction which data most spreads out. In other words, it has the largest possible variance. The other principal components have the highest variance with the constraint that they should be orthogonal to all other previous components. These principal components can be obtained using the covariance approach or the singular decomposition method. Since the covariance method is more straightforward, I explain the method of achieving the principal components using its intuitive steps.

Consider a matrix, X, with rows being the observations and columns being the variables. PCA maps each row vector to a new space with principal components being its basis. For example, one can extract the first two principal components of a high dimensional dataset and map these high dimensional data points to a new space with only two dimensions being the principal components. This can be a very useful preprocessing step. To extract the principal components, first of all, the empirical mean is removed from the columns. Then, the eigenvectors of the covariance matrix are sorted in terms of their eigenvalues. The first eigenvector which corresponds to the largest eigenvalue would be the principal component.

2.2. Independent component analysis

Independent component analysis is a generative blind source separation model utilized for extracting source signals from a linear mixture of signals without having a priori knowledge about the nature of mixture. ICA works on three assumptions that need to be considered carefully before applying it on a problem. The first assumption is that the source signals are uncorrelated and statistically independent. In other words, the value of a signal at a specific time does not tell anything about the value of other signals at that time. The second assumption is that the source signals are mixed linearly. The last assumption is that the source signals should not follow Gaussian distribution.

In order to clarify on ICA, I refer to a concrete example of the classic 'cocktail party problem'. Imagine a cocktail party is held in a room where four people are speaking simultaneously, while their sound is recorded by four microphones. As the microphones are positioned in different locations, the individuals' speech signal will contribute differently to each microphone. Furthermore, each microphone records the sound of all four individuals. This problem can be written using algebraic notations.

$$x_1 = a_{11}s_1 + a_{12}s_2 + a_{13}s_3 + a_{14}s_4$$

$$x_2 = a_{21}s_1 + a_{22}s_2 + a_{23}s_3 + a_{24}s_4$$

$$x_3 = a_{31}s_1 + a_{32}s_2 + a_{33}s_3 + a_{34}s_4$$

$$x_4 = a_{41}s_1 + a_{42}s_2 + a_{43}s_3 + a_{44}s_4$$

x_i denotes the output of the i-th microphone; s_i represents the i-th individual's speech signal; and, a_{ij} stands for a weight which depends on the distance between i-th microphone and j-th person. This problem can be further simplified using matrix notation:

$$x = As \tag{1}$$
$$u = Wx \tag{2}$$

In (1), s denotes the source signals; A represents the mixing matrix; and, x stands for the microphone output. If we knew A, we could apply straightforward linear Algebra methods to extract s; however, we have no clue of how the source signals are mixed in practice. In other words, there are two unknowns and one known variable. In this case, one can apply ICA because its assumptions hold; the amplitude of each voice at a specific time does not tell us anything about the amplitude of another voice at that time, so the independence assumption of the source signals holds. The source signals are mixed linearly, and they do not follow Gaussian distribution. There are well-known algorithms which find an estimation u of the source signal s by computing the separating matrix W.

Infomax [1,2] and FastICA[3,4] are the two most popular algorithm for ICA. Infomax finds the separating matrix W by utilizing negentropy and minimizing the mutual information of the estimated source signals u_i. On the other hand, FastICA exploits kurtosis and maximizes the non-Gaussianity of the estimated u_i. FastICA is based on central limit theorem. An observation in central limit theorem tells us that the distribution of two independent random variables is closer to Gaussian than the two original variables. Furthermore, Hyvärinen and Oja [3] has demonstrated that maximizing the non-Gaussianity of the estimated source signals can be reduced to minimizing their mutual information. He also concludes that the source signals must be non-Gaussian; otherwise, ICA cannot be applied.

Both Infomax and FastICA implement centering, whitening, and dimensionality reduction as the preprocessing steps. These steps have been demonstrated using a simple but intuitive example in

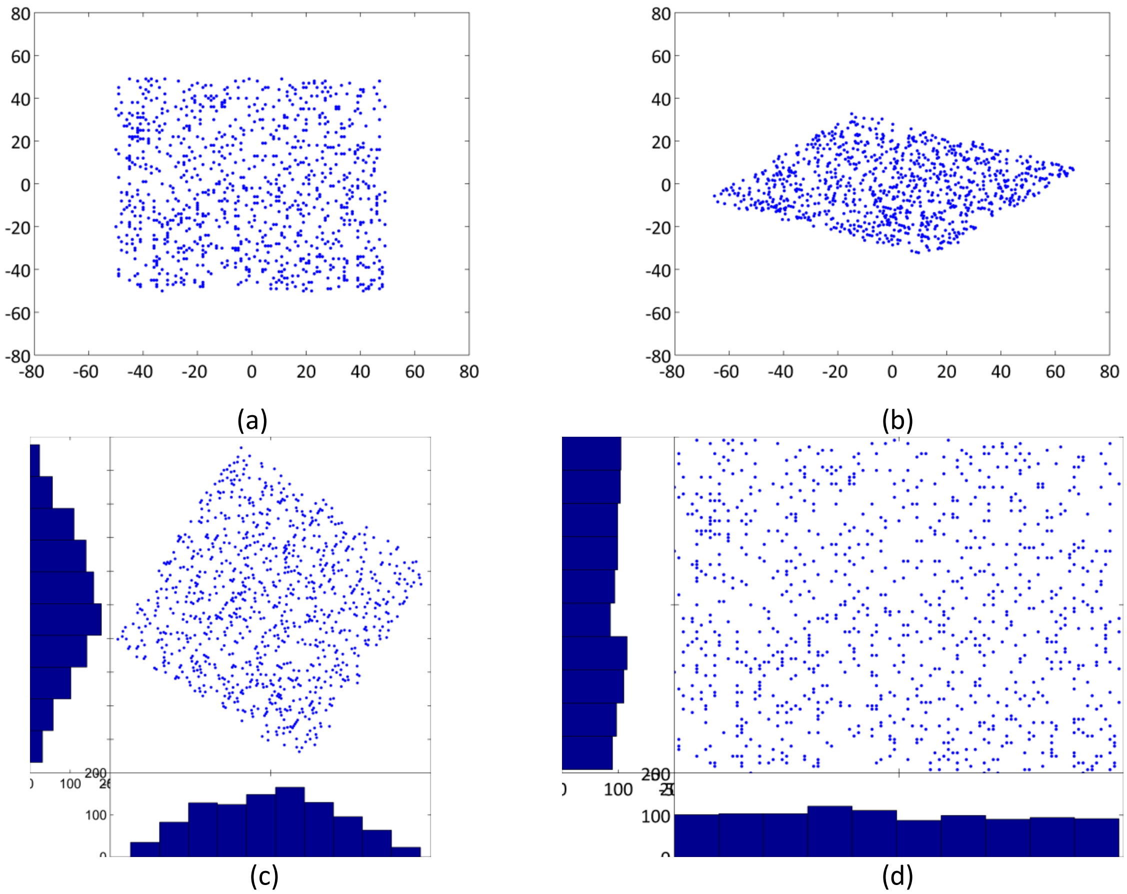


Figure 2: (a) shows the original data generated from a uniform distribution. The two source signals are defined by horizontal and vertical coordinates of the data points. (b) demonstrates how the source signals are linearly mixed. (c) depicts the whitening step. (d) reveals the estimated source signals found by FastICA.

Figure 2.a demonstrates the original data. 1000 data points are generated from the uniform distribution. The first source signal s_1 is defined by the horizontal coordinate of the data points, and the second source signal s_2 is determined by the vertical coordinate of the data points. The two source signals are generated independently, and they follow uniform distribution. Hence, this is a perfect example for ICA. Figure 2.b exhibits linear mixture of the two source signals. (3) and (4) shows how they are mixed.

$$\begin{cases} x_1 = 0.54s_1 + 0.84s_2 \ (3) \\ x_2 = 0.42s_1 + 0.27s_2 \ (4) \end{cases}$$

Figure 2.c shows the whitening step. This step is implemented by removing the mean from the data and multiplying it by the inverse of square root of the covariance matrix; it is worth mentioning that whitening has restored the original shape of the data, and ICA just needs to rotate it. The output of FastICA algorithm is also manifested in Figure 2.d; a geometric interpretation of the ICA is that it rotates the axis and minimizes the Gaussianity of data projected onto each axis.

ICA has been applied in various fields including but not limited to audio signal processing [5-7], image processing [7-12], bioinformatics [8-26], time series and financial data [27-30], and text document analysis [31-33]. WordICA is another successful application of ICA in natural language processing.

2.3. WordICA

WordICA [36] is an unsupervised machine learning approach that can automatically find word features from unannotated corpora. These word features can be utilized for tasks such as word clustering. Since WordICA is an unsupervised technique, one does not have to worry about annotating a large corpus, which makes it an efficient method. In fact, sufficient corpus for training classical language technologies is not available in many languages including but not limited to Farsi. In addition, these features can be exploited to produce automatic lexical resources; these resources will be useful in developing applications such as natural language interface and machine translation systems.

ICA is a numeric algorithm, so it is required to transforms words into numbers. One way to achieve this transformation is applying the bag of words model. In this model, a word is represented by a vector with one element equal to one and other elements equal to zero. Thus a word is represented by a vector in a numerical space whose dimension is determined by the number of context words. In addition, it should be emphasized that the word orders and their dependencies are ignored in this model. Besides that, the dimension of this space can be reduced by utilizing SVD, PCA, or other dimensionality reduction methods.

Word-word matrix and word-document matrix are the two conventional ways of transforming a text document into a numerical representation. The rows of word-word matrix indicate the analyzed words, and its columns indicate the context words. The analyzed words can be the set of the N most frequent words or any other set of N words. The context words come from the set of M most frequent words. This will form an N-by-M matrix denoted by X whose element x_{cn} means the frequency of c-th word occurred with a specific distance from the n-th context word. This distance can be, for example, the two immediate preceding and two immediate following words. In addition, in word-document matrix, rows represent the N most frequent words, and columns denote the documents. Each element of word-document matrix refers to the frequency of a word in a document.

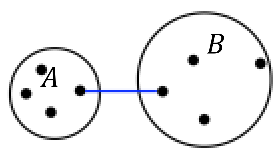
The rows of word-word or word-document matrix represent a vector in a high dimensional space. Because these matrices are sparse, it is safe to express that the components of these vectors follow a distribution which is far from Gaussian distribution. One can also view the rows of these matrices as mixture signals. In other words, some underlying source signals have been mixed and formed them. As the non-Gaussianity assumption is valid here, we can apply ICA to extract the underlying factors. These underlying source signals are in the same space as the words; thus, they can be exploited for word clustering.

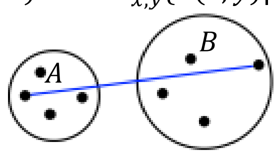
Word clustering using ICA can be achieved by comparing word vectors against independent component vectors extracted from word-word or word-document matrix. Some distance measure such as cosine distance is employed to determine the distance between a word vector and a component vector. The closest words to a component can form a cluster. One interpretation of such cluster is that its words have similar document or context distribution.

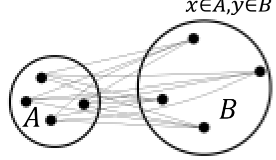
2.4. Hierarchical clustering

Hierarchical clustering [126] is a well-known clustering method by which one can build a hierarchy of clusters. This clustering method can be either agglomerative or divisive. In agglomerative hierarchical clustering, one starts with each data points being placed at its own singleton cluster; then, two closest clusters are merged iteratively until all the data points are merged into one single cluster. Thus, agglomerative clustering is a bottom-up approach. On the other hand, divisive clustering is a top-down approach; in other words, all the data points are at one single cluster at first; then, iteratively, the clusters are split into smaller clusters until every data point is at its own cluster.

In hierarchical clustering, one does not need to know the number of clusters in advance. It just requires distances between data points and a measure of similarity between clusters. The distance between every pair of data points can be measured using Euclidean, Cosine, or any other metrics. However, computing the distance between two clusters is trickier. There are seven methods with which one can measure the similarity between two clusters.

1. Single linkage: the distance between two clusters A and B is the shortest distance between any object in A and any object in B.
$$D(A,B) = min_{x,y}\{d(x,y)|\ x \in A, y\ \in B\}$$


2. Complete linkage: the distance between two clusters A and B is the largest distance between any object in A and any object in B.
$$D(A,B) = max_{x,y}\{d(x,y)|\ x \in A, y \in B\}$$


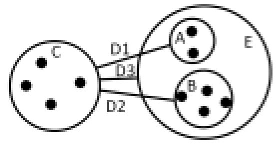
3. Group average distance: the distance between two clusters A and B is the average distance between any object in A and any object in B.`this method is also known as Unweighted Pair Group Method with Arithmetic Mean (UPGMA).
$$D(A,B) = \frac{1}{|A| \times |B|} \sum_{x \in A, y \in B} d(x,y)$$


4. McQuitty's method: suppose that there are three clusters with labels A, B, and C. Also, imagine that clusters A and B are merged together to form a new cluster E. Then, the distance between

cluster E and C is computed by a weighted average of the distances between C and sub-clusters of E which are A and B. this method is also known as Weighted Pair Group Method with Arithmetic Mean (WPGMA).

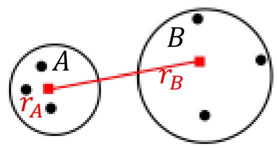
$$D(E,C) = \frac{(|A| \times D(A,C) + |B| \times D(B,C))}{(|A| + |B|)}$$

$$D(E,C) = \frac{2 \times D_1 + 4 \times D_2}{2+4}$$



5. Centroid: the distance between two clusters A and B is the distance between their centroids $r_A$and r_B.

$$D(A,B) = d(r_A, r_B)$$



6. Median: the distance between two clusters A and B is the Euclidean distance between their weighted centroids $\bar{r}_A$ and $\bar{r}_B$, where the weighted centroids are defined recursively. For example, if A was created from clusters p and q, then its weighted centroid is defined in this way:

$$\bar{r}_A = \frac{1}{2}(\bar{r}_p + \bar{r}_q)$$

$$D(A,B) = ||\bar{r}_A - \bar{r}_B||_2$$

7. Ward: the distance between cluster A and B is computed in this way; first, for each cluster we compute the sum of squared deviations from the cluster's centroid r_A and r_B. Then we merge these two clusters and compute the sum of squared deviations from the newly created cluster's centroid r_{AB}. Finally, we sum up all the first two values and subtract the third value from it; finally, we take a weighted average of it.

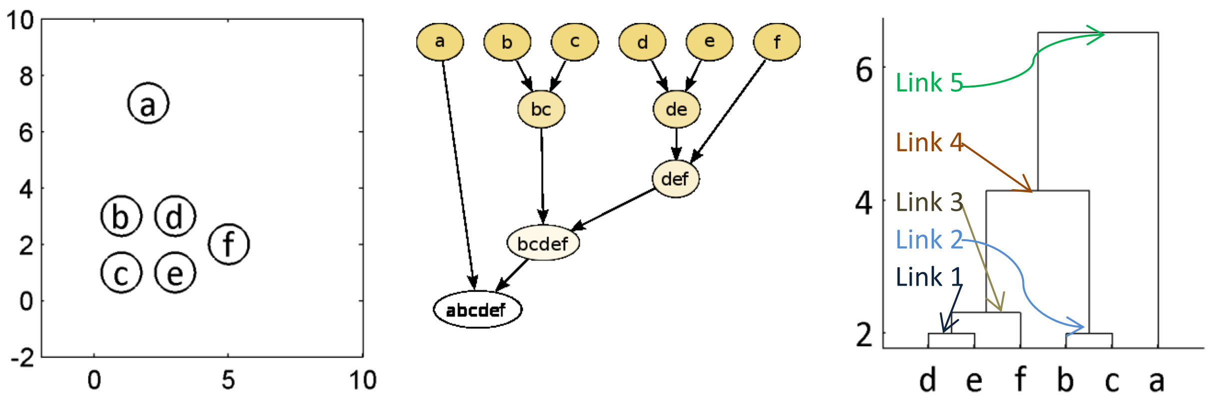
$$D(A,B) = \frac{|A| \times |B|}{|A| + |B|} \left| \sum_{x \in A} (x - r_A)^2 + \sum_{x \in B} (x - r_B)^2 + \sum_{x \in AB} (x - r_{AB})^2 \right|$$

One can select any of the above method to compute the distances between two clusters. However, we should notice that the choice of distances between two objects limit our options. For example, ward, centroid, and the median method are appropriate only when the distance between two objects is Euclidean.

The result of a hierarchical clustering is visualized using a dendrogram. For example, Figure 3.a demonstrates six 2-dimensional data points. Euclidean metric is applied to determine the distance among data points; at first, each data point is placed at its own single cluster; then, iteratively, the two closest clusters are merged using ‘average’ method. As you can see in the above figure, every two objects will be merged at some level. The height corresponds to this level is called ‘cophenetic distance’ [39].

2.4.1. Cophenetic correlation coeffiecients

There are two distinct types of distances among objects. The first type of distance is computed by considering the vectors that define the objects in the original vector space model. It is referred to as the original distance. Let's consider the six objects in Figure 3.a; each one of them is defined by a two-dimensional vector. For example the numerical representation for object b=(1,3) and for object c=(1,1). If Euclidean metric is selected, the distance between object b and c is two.



(a) (b) (c)

Figure 3: (a) represents the data points which are going to be clustered. (b) displays a traditional hierarchical clustering tree. (c) denotes the resulting dendrogram. The distance between data points is measured using Euclidean metric, and 'average' method is utilized for computing the similarity of clusters.

The second type of the distance is read from the height of the links in a dendrogram. A link in a dendrogram is a straight line connecting two branches. The height of a link determines the distance among the objects that fall into the left-side branch and the objects which lie in the right-side branch. This distance is called 'cophenetic distance' [39]. For instance, the cophenetic distance between object b and c in Figure 3.c is 2. In this case, the cophenetic distance correctly reflects the original distance which was also 2. On the other hand, the cophenetic distance between object b and d in Figure 3.c is 4.13, while the Euclidean distance between these two objects is 2.

Cophenetic correlation coefficient measures how faithfully a dendrogram preserves the original pairwise distances among data points.

$$c = \frac{\sum_{i<j}(x(i,j)-\bar{x})(t(i,j)-\bar{t})}{\sqrt{\left[\sum_{i<j}(x(i,j)-\bar{x})^2\right]\left[\sum_{i<j}(t(i,j)-\bar{t})^2\right]}}$$

Where $x(i,j)$ is the original distance between object i and j, while $t(i,j)$ represents their cophenetic distance.

2.4.2. Extracting clusters from a dendrogram

The height of a link illustrates the cophenetic distance between the objects in the left-side and objects in the right-side branch. Cophenetic distance is an approximation for the original distances among the objects. In a densely packed area, the height of a link is compatible with the height of the links below it, which means the cophenetic distances in a dense area are consistent with each other. Thus, the objects in a dense area are not far from each other, while they are farther from the other objects. Hence, a dense area is an indication of a single cluster. Besides that, one can decide about the number of underlying clusters by observing a dendrogram. This makes the dendrogram a very simple but powerful information visuzalition tool using which one can observe the potential clusters.

2.4.2.1. Extracting arbitrary number of clusters from a dendrogram

After deciding about the number of clusters, one can cut the denrogram to partition data into clusters of objects. For example, the objects in Figure 3.c can be divided into three clusters; three clusters are selected because there are three dense areas. To implement the partitioning, imagine a hypothetical horizontal line. If this hypothetical line is laid over the highest link, all the objects will be put into one cluster. If the line is placed right below the highest link and above the second highest link, then the data is partitioned into two clusters. Similarly, if the line is situated right below the second link, the data is divided into 3 clusters. In other words, to divide the data into 3 clusters, one should cut a dendrogram through three branches. The objects below each branch will be deposited into a single cluster. For instance, when the dendrogram in Figure 3.c is cut into three partitions, these clusters are extracted: {d,e,f}, {b,c}, and {a}.

2.4.2.2. Extracting clusters from a dendrogram based on the inconsistency coefficient

In the method that has just been expounded, the task of dividing data into distinct clusters was based on the observation of densely packed area in a dendrogram. This observation was itself built upon the idea of inconsistent links. The inconsistency of a link with respect to the links below it can also be measured mathematically [40, 41]. One can apply the inconsistency coefficient formula to map each link to a real number. This number will give us a clue of how much a link complies with the average height of the links below it. Abrupt change in the inconsistency coefficient of a link reveals that the height of that link does not agree with height of the links below it. This delineates a potential natural division in the data.

The higher the inconsistency of a link, the less similar are the objects connected by that link. In other words, although the objects placed below the left branch of the link might be close to each other, they are farther from the objects situated below the right branch. Thus, an inconsistent link illustrates a border of cluster or a natural division among data. In order o find the inconsistent link, one must find the inconsistency coefficient of all the links in a dendrogram. Then using a cutoff value, one can cut through the dendrogram. Table 3, demonstrates the influence of the cutoff value on the clustering of objects visualized in Figure 3.a.

Table 3: Clustering the 2-dimensional objects using the inconsistency coeeficient

Link id	Inconsistency coefficent		Cluster labels		
		Objects	Cutoff=1.6	Cutoff=1.5	Cutoff=0.8
		a	1	1	2
1	0	b	1	2	1
2	0	c	1	2	1
3	0.7071	d	1	2	3
4	1.4847	e	1	2	3
5	1.5937	f	1	2	3

Agglomerative hierarchical clustering was applied to six objects which you can observe their coordinates in Figure 3.a; the dendrogram demonstrated in Figure 3.c is the result of such hierarchial clustering. A dendrogram shed light on the number of possible clusters, but what we still need to extract is the clustering label of each object. In order to determine to which cluster each object belongs, we must either cut the dendrogram into horizontal slices or cut it using the inconsistency coefficient of the links. Here, I clarify on the second method using a concrete example.

In Table 3, you can see the inconsistency coeeficient of every five links in the dendrogram of Figure 3.c. The inconsistency coefficients of link 1 and 2 are zero because their children are leaf nodes. The other three links have nonzero inconsistency coefficient. In terms of the inconsistency coefficient value, link 3 is more consistent with its below links than link 4 and 5. This is also consipicuous in the dendrogram; the height of link 3 does not change significantly from the height of link 1, while the height of link 4 and 5 vary much more from the height of their below links. Hence, the inconsistency coefficient of link 3 indicates a border in the data. Its value can be assigned to the cutoff variable. The value of this variable will determine the clustering labels.

If cutoff is set to 1.6, all the objects will be distributed into one cluster since the inconsistency of all the links are smaller than the cutoff value. If cutoff is set to 1.5, the data would be divided into two partitions. That's because the inconsistency coefficient of link 5 is higher than cutoff=1.5, so that link will be cut through; all the objects that are below the link 5 go into one cluster, while the objects below the right branch go into the second cluster.

The process of dividing a dendrogram based on the cutoff value is implemented in top-down approach. For example, when cutoff is 0.8, link 5 is cut at first; the objects below its right branch, namely {a}, go into the first cluster. Then, link 4 is cut; the objects below its right branch, namely {b, c}, put into the second cluster, and the objects below its left branch, namely {d, e, f}, placed in the third cluster. As the inconsistency coefficient of all the remaining links are below the cutoff value, the partitioning process is stop at this stage leading to the extraction of three clusters.

As you have seen, using any threshold for the cutoff variable in this example results in a horizontal division of the dendrogram; however, one should notice that the clustering based on cutoff value does not necessarily map to horizontal slices of a dendrogram. As a result, dividing a dendrogram according to thresholding of the inconsistent links is less intuitive than the horizontal division of the dendrogram.

2.4.3. Applying hierarchical clustering to one-dimensional distributions

One can also apply hierarchical clustering to 1-dimensional distributions. For example, in Figure 4.a, you can see 6 normal distributions. Each one of the distributions has been represented by a vector of 500 components. In other words, one can also view each distribution as a point in a 500-dimensional space. In this case, cosine distance can be applied to determine the similarity among distributions. Furthermore, one can also apply hierarchical clustering to put the distrubutions in hierarchical clusters.

As you can see in Figure 4.b, at first, cyan and yellow have been merged; this is simply because the mean of the two distributions are almost identical. Next, the magenta and the red distributions are merged to it. Blue and green distributions have been merged in the last iteration. The height of the link at which two clusters are merged can show inconsistency if it is higher than the height of the links below it. Thus, one interpretation of Figure 4.b is that {yellow, cyan, magenta, and red} are in one cluster, while blue and green are in their own separate clusters.

2.4.1. Verify the clustering tree

In order to verify a hierarchical clustering, one should measure the similarity among objects using the cophenetic distances and check how well these cophenetic distances reflect the original distances among objects. This task can be analyzed using cophenetic function that is provided by popular statistical toolboxes. This function returns a value which is called ‘cophenetic correlation coefficient’ [39]. The closer this value is to 1, the more precisely the clustering solution reflects the natural divisions in data.

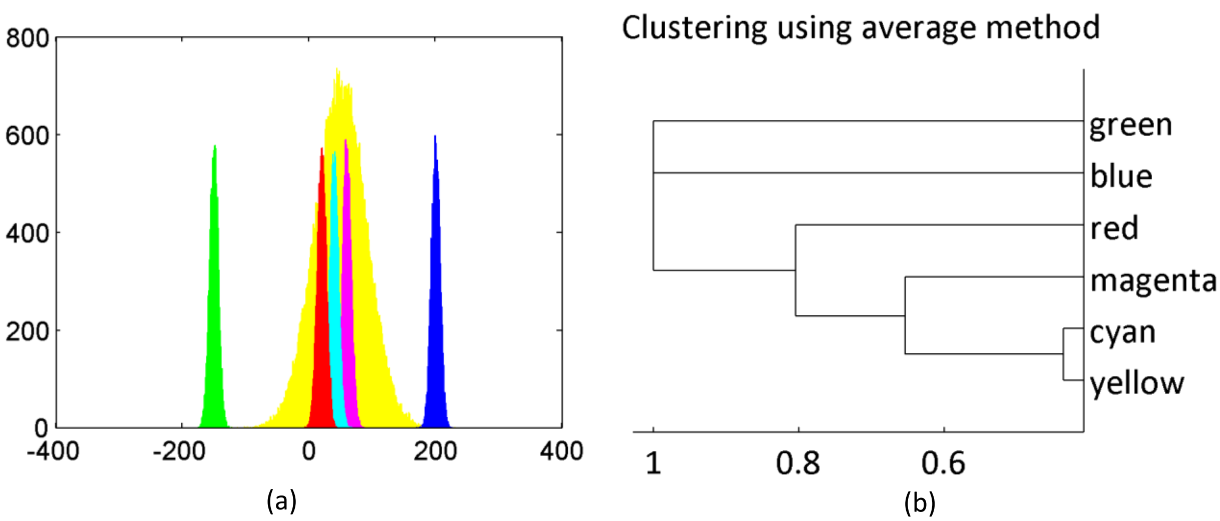


Figure 4: (a) represents 6 normal distributions; the values of each one of these distributions have been stored in a vector of 500 components. (b) shows the hierarchical clustering of the distributions. Average method and cosine distance have been applied.

2.5. Motion data generation and the web-based questionnaire

In order to create motions, Arena made by NaturalPoint OptiTrack was used as the motion tracking software system equipped with 24 infrared cameras working at 100 frames per second [141]. The cameras track retro-reflective markers attached to an actor. Furthermore, the mocap software calculates the 3D coordinates of the markers and also solves the joint rotations for a skeleton in each frame. Every frame is stored in a file with BVH format (BioVision Hierarchy). Such raw motion data allows animating a humanoid character with a segment of a motion and saving the animation to a video file.

Although the raw mocap data is a good representation for a single frame of a motion, its high dimensionality characteristic makes it an inappropriate representation for the whole motion which can have thousands of frames. Therefore, based on the raw motion data, higher level features (602-dimensional) for representing motion segments were calculated that included the means and standard deviations of coordinates, velocities, and the accelerations of different body parts. See [142] for a more detailed explanation of the contents of this 602-dimensional dataset.

Each one of the two actors performed 12 locomotions with varying styles such as ‘sad’, ‘slow’, ‘regular’, ‘fast’ and ‘angry’. From these 24 motions, 100 additional motions were created by time warping and interpolating the raw motion data, which allowed variation in both verbs and modifiers. To make the interpolation task easier, the actors start all their movements with their right legs from the same position and toward the same direction. See chapter 3 from [142] for a detailed description of how these motions were generated and interpolated.

A web-based questionnaire has been designed for collecting annotations. Annotators can describe a motion video with one verb and from zero up to three modifiers. In total, there are 124 videos which have been divided into 3 sets namely A, B, and C. Set A contains the 24 unmodified videos. Set B has 40 videos with 50%-50% interpolations; and the rest of videos were included in set C. Every motion was presented with a stick figure character as shown in Figure 5.

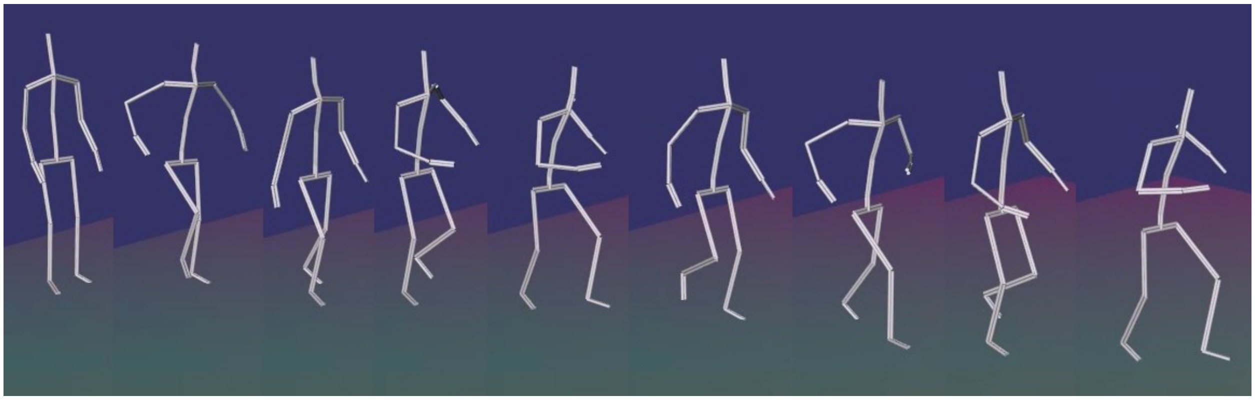


Figure 5: a stick figure character which is walking angrily.

2.6. Finding synonyms of motion-related words using video indexes

One way of finding the synonyms of the annotated words is to manually categorize them by putting them into some groups. In this case, each group consists of all synonyms and the different ways of saying the same words. For example, a group can be the synonyms of 'slowly'. Some people might use an adjective like 'slow'; some might use an adverb like 'slowly'; and, some might even write a word with an incorrect spelling. Besides that, some people might have used synonyms of 'slowly' such as ‘leisurely’. In this approach, after observing the resulting lexicon, the words were grouped manually. Now, we are interested in finding these categories automatically. These categories can be extracted using two approaches. In the first approach, words are defined in the context of the videos in which they emerge; then, using the video ID as the contextual information, an automatic thesaurus of the motion annotated words is extracted. We refer to this approach as the **indexical grounding**. In the second approach, in addition to utilizing the video indexes as the context of words, the vectors corresponding the annotated motion words are normalized using the motion data. In this case, we would expect more accurate synonyms to be extracted. Otherwise, all our effort of employing motion data, coming from the visual modality for symbol grounding, in the construction of a thesaurus would rise up in the smoke. We refer to this case as the pattern grounding.

I clarify these two approaches using a step by step description. First, we represent a word by a vector. The dimension of this vector is 124; there is one dimension for each video. Let's say, we are going to form a vector for the word 'slowly'. The first component of this vector represents the frequency of 'slowly' in the first video. In other words, we count how many people have described the first video as 'slowly'. Then, its second component will be the frequency of ‘slowly’ in the second video. Other components of this vector are defined in the same way. Finally, we will have a vector which represents the word 'slowly' in a 124-dimensional space. We can do the same process for all other words, which leads to a matrix whose rows represent words and columns represent videos. One can find synonyms of all words by applying a distance metric on this matrix. It is worth mentioning that in this approach, the only anchor which connects the annotated words to the physical motions is the video index.

Our observations show that adverbs are distributed in different places of the vector space. This is due to the fact that one can describe different verbs with the same adverb. For example, one can describe 'running', 'walking', and 'limping' by 'slowly'. Hence, adverbs can be used for many videos; this means that adverbs are distributed evenly which is closer to uniform distribution. Because of this observation, we form such matrix only for verb and verb-adverb combination.

2.6.1. Finding synonyms of verbs automatically

We can extract categories of verbs automatically from the verb-video matrix. The rows of this matrix represent verbs, and the columns represent videos. Since each verb is represented by a vector, one way to find its synonyms is by computing the distances between vectors. One can use either Euclidean or cosine distance measure for computing the similarity. The closer two vectors, the more similar their corresponding verbs are.

2.6.2. Finding verb-modifier synonyms automatically

Similarly, one can represent a verb-adverb with a vector. I have to say that by modifier, I mean both adverbs and adjectives, and since most of the annotated modifers are adverbs, modifiers and adverbs are used interchangeably. We can count the frequency of each verb-adverb in all videos and put the resulting numbers in the components of a vector. For example, a vector can represent 'walk-slowly'. The first element of this vector shows the frequency of 'walk-slowly' in the first video; in other words, it tells us how many people have described the first video by 'walk-slowly'. After each verb-adverb is represented by a vector, one can compare the distance between vectors and extract similar verb-adverbs.

2.7. Finding word similarities by fusing visual features

In the previous section, similar words have been extracted from word-video matrix, where we only used the video index as the only anchor to connect the words to the physical motions; column 1 of the word-video matrix denotes the video whose index is 1. Column 2 represented the second video, and other columns were defined in the same way. In other words, in the previous section, we considered the distribution of words in a vector space in which the video distributions is neglected. By video distribution, I mean the similarity of videos. One can compute the similarity of videos by extracting some feature vector from motion videos. Thus, each motion video is denoted by a vector. The components of such vector denote data such as the speed of limb ends and their direction. This feature vector has been calculated using the software by which the motions were created.

By comparing the corresponding vectors of two motion videos, one can compute how similar they are. This comparison can be done in the original vector space or in a reduced-dimensional vector space. In this experiment, we have computed the video similarities in the original space because the accuracy is more important to us than the resulting time-complexity. Finally, the word-video matrix is normalized using the video-similarity information as clarified in the next section. We refer to this normalization as the **pattern grounding**.

2.7.1. Normalizing word-video matrices using video-similarity

In order to compute their similarities, videos have been transformed to a vector space so that a numerical representation can be extracted. To achieve this numerical representation, 602 features are obtained from the motion capture system. Coordinates of limb ends, their direction, and their speed are examples of such features we have extracted from each motion video. Thus each one of 124 videos is denoted by a vector in a 602-dimensional space. Now that we have the raw data, we can compute the similarities of the videos. This can be achieved by computing the distances among vectors by applying a similarity measure such as cosine distance.

The next step in fusing visual features with the textual features is to normalize both verb-video and verb_adverb-video matrix based on the motion data. If a video is very close to another video in the 602-dimensional vector space, one can expect their verb and adverbs to be similar. For example, suppose video i and video j are very close, so we expect that column i and column j of the verb-video and verb_adverb-video to be similar. If not, some people might have annotated these two similar videos by synonym verbs. Besides that, the annotators might have forgotten to use all the possible adverbs, or it might be caused by the limited option they had as people can annotate a motion video with at most 3 modifiers.

One way to normalize word-video matrices is to define a neighborhood for every video. This neighborhood can be determined by a fixed distance or set differently for each video. We have decided to start symbol grounding experience with a fixed neighborhood distance as it is simpler to implement. The distances of the closest neighbor of every video have been computed; the longest of such distances is 0.2575, and the shortest one is 0.0119. In other words, there is a video whose closest neighbor is 0.2575 far away. This has propelled me to select 0.3 as the fixed cosine distance that determines the neighborhood of a video. In this way, every video has at least one neighbor. Figure 6 shed light on the distribution of the number of neighbors.

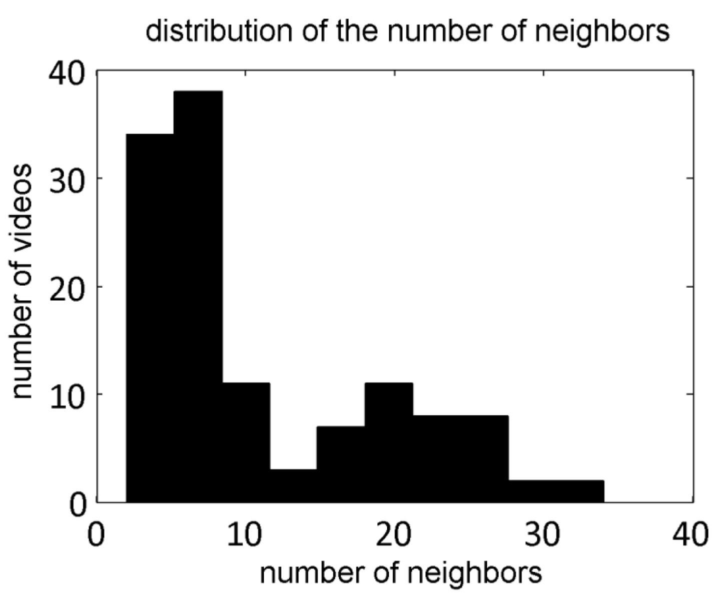


Figure 6: every video has at least 1 video in its neighborhood. In addition, the majority of the videos have at most 10 videos in the neighborhood.

The next step of symbol grounding is to determine how much influence a video has on its neighbors. Suppose video j and video k are two neighbors of video i. Then, column j and column k of both verb-video and verb_adverb-video matrix should be normalized by columns i of the aforementioned matrix. This normalization can be achieved by adding a multiple of column i to column j and k. Now, we should determine what multiple of a column should be added to its neighbors. The steps of normalization are further clarified in the following algorithm.

An algorithm for normalizing verb-video and verb_adverb-video matrix by incorporating the motion data

Input: α is the maximum addition proportion which can be a positive value between 0 to 1

```
for every  video i
        neighborhood= extract neighbors from the motion feature space
        sort the neighbors according to their distance from video i
        shortest_distance=the distance of the closest neighbor from video i
        for every video j in the neighborhood of video i
            current_distance= distance of video j from video i
            critical_value=(shortest_distance / current_distance) × α
            add (critical_value × column i) to column j of word-video matrix
        end
end
```

Using the above algorithm, the neighbors of a video are normalized by the video. Such normalization is implemented by adding a multiple of the video vector to its neighboring video vectors. This addition is carried out with respect to the distance of a neighbor to the video. The closer a neighbor is to the video, a larger multiple of that video is added to it. This is controlled by the critical_value. In addition, α determines the maximum addition proportion. For example, if α=0.5, half of the video vector is added to its closest neighbor.

The above algorithm has been applied to both verb-video matrix and verb_adverb-video matrix. Since the motion data originates from the visual modality, and it was fused with the textual features, we have implemented symbol grounding. Next, we should analyze whether this kind of normalization and hence symbol grounding is effective or not.

2.7.2. Extracting synonyms from the normalized word-video matrix

We normalize verb-video matrix using the aforementioned algorithm. Therefore, the symbols, which in this case are the verbs, are grounded using visual data. Next, we compute vector distance among all the rows of this normalized verb-video matrix and select five closest synonyms of every verb.

The verb_adverb-video matrix is also normalized in the same way as the verb-video matrix was normalized. First, the videos should be transformed into numerical vectors. This has been achieved by utilizing motion data. After that, the distance between every pair of videos is computed and video distance matrix is built. Next, a fixed neighborhood distance is selected; using such distance, the neighbors of every video are determined. Finally, we apply the normalization algorithm that was mentioned in the previous section, to the verb_adverb-video matrix.

2.8. Automatic translation

One can utilize videos as the contextual information and translate verbs and verb-adverbs from a language to another language. Furthermore, we want to observe if incorporating motion data can improve the quality of translation. Since this motion data is extracted from the visual modalities, this kind of incorporating motion data with text data is considered to be symbol grounding.

2.8.1. Automatic translation using indexical grounding

In this section we want to check how good translating verbs and verb-adverbs is if we use only the contextual information which is video indexes. We employ videos as the contextual information to transform verbs into vectors. One can do this by counting the occurrences of each verb in every video. Since there are 124 videos, the resulting vector will be 124-dimensional. The first component of such vector represents the number of occurrences of the corresponding verb in the first video. The other components of the vectors are defined in the same way.

Now that we have a vector representation for every verb, we can compare the verb vectors of two languages by exploiting a distance measure such as 'cosine distance'. The next step is to select, for example, 5 closest verbs in the target language for each verb in the source language, and count how many of these verbs are semantically related to the verb in the source language. Of course, the higher this number the better the translation is.

Frequency of all possible combinations of verbs and adverbs in each video are also extracted. Using this frequency information, one can convert a verb-adverb into a 124-dimensional vector. We put these vectors into the rows of a matrix so that each row of the matrix represents a verb-adverb combination, and each column represents a video. We execute this process for verb-adverbs of both the target and source language. Finally, the distance matrix is created by computing the proximity between each row of the two matrices. After that, for each verb-adverb, we sort the closest vectors; the corresponding verbs-adverbs would be good candidates for translation.

2.8.2. Automatic translation using pattern grounding

In this section, we translate the verbs and verb-adverbs by grounding them using the motion data coming from the visual modality. The verb-video matrix is created for the annotated words in both the target and source language. Then, by converting videos into vectors using the motion data, one can determine the closest neighbors of each video. Finally, the verb-video matrix is normalized using the normalization algorithm mentioned previously.

Frequency of all possible combinations of verbs and adverbs in each video are extracted. Using this frequency information, one can convert a verb-adverb into a 124-dimensional vector. We put these vectors into the rows of a matrix so that each row of the matrix represents a verb-adverb combination, and each column represents a video. We execute this process for verb-adverbs of both the target and source language. Then, we normalize these two matrices using the normalization algorithm. Finally, a distance matrix is created by computing the proximity between each row of the two matrices. After that, for each verb-adverb, we sort the closest vectors; the corresponding verbs-adverbs would be good candidates for translation.

3. Experiments and the results
For the first part of the project, word features have been extracted from two available Farsi corpora. For these corpora, both word-context and word-document matrices have been formed, and then ICA has been applied to both of these matrices. The ICA yielded some word features which were exploited in the detection of automatic word categories. Next, these categories are labeled manually. The datasets, the preprocessing step, and their word clusters are described in section 3.1.

The second part of the project is about the utilization of the motion data for the purpose of symbol grounding. At first, we made some observations about the potential relationship between the textual and visual features. The videos were mapped to 2-dimensional vectors using PCA and visualized by pies. After that, the verbs and verb-adverbs, emerged on videos, have been presented on the resulting pies. This revelatory observation propelled us to fuse the features from two distinct modalities of text and vision. The annotated verbs and verb_adverbs have been represented by vectors; these vectors were normalized using the motion data extracted from the visual modality. Such normalization turned out to be fruitful in increasing the accuracy of synonym detection and translation processes.

3.1. Word clustering experiment
Automatic word clusters have been extracted from two Farsi corpora. Independent component analysis (ICA) has been applied on these two corpora. ICA does not require any apriori knowledge or assumption about the language; thus, no preprocessing is necessary; however, some preprocessing tasks such as removing the plural sign can be useful.

ICA has been employed to extract 10 components from the word-word and word-document matrix. An extracted component can be perceived as an abstract embodiment of a word cluster. In addition, since both the extracted component and the word vectors are defined in the same vector space, a similarity measure such as cosine distance can be used to determine which words go with which cluster. In other words, the proximity between an extracted component and the word vectors lead to automatic construction of a thesaurus.

3.1.1. Datasets
The first dataset contains 10010 news articles from Alef which is a popular online news agency in Iran. The second corpus includes 20872 news documents from Tabnak which is another popular online news agency in Iran. The documents of both corpora are concatenated into two big files [128]. In addition to the content of each document, extra information such as the title and category is also incorporated into the file. In the preprocessing step, meaningless symbols such as semicolons, one-letter and two-letter words have been removed. Furthermore, a few inflectional morphemes such as (plural sign=ها) and (present or past continuous tense sign=مي) have been excluded.

3.1.2. Extracting features from word-document matrix of Alef dataset
In the first experiment with Alef dataset, word-document matrix is formed. 500 words ranked from 201 to 700, in terms of frequency, make up the dictionary, so the 200 most frequent words have been ignored. In other words, the word-document matrix associated with the Alef dataset has 500 rows and 10010 columns. Thus, each document is defined in a 500 dimensional space. We can also view the words in a 10010 dimensional space. Since we want to cluster words, we work with rows and extract 10 components using FastICA package developed in Helsinki University of Technology by Hurri and his colleagues [37]. Each component is an abstract representation of a word cluster in a 10010 dimensional

space. Then, cosine distance is computed between each one of these five hundred words and each component. In other words, a distance matrix is created. The rows of this distance matrix denote the component and the columns represent the words, so in this case, it is a 10 by 500 matrix. For each component, five closest words are selected; this can be done by sorting the rows. In addition, the label for each cluster is assigned by considering a relation between the meanings of the words; which means that the label assignment is not automatic.

Table 4: Automatic thesaurus construction using word-document matrix extracted from the Alef corpus	
Detected category	Extracted words
economy	اصل(principle, origin) خصوصی(private) توزیع(distribution) اقتصاد(economy) سازی (to make; it is used together with another term; in other words it is a half token)
politic	عراق(Iraq) آمریکایی(American) شکل(shape, form) حرکت(movement) همکاری(collaboration)
management	شرکت(company) دولتی(government-owned) کاهش(reduction) تعداد(number) مدیریت(management)
politic	غزه(Gaza) ماده(material, bill=law proposal) صهیونیستی(zionist) زنان(women) آمریکا(USA)
economy	مصرف(consumption) رشد(increase) سرمایه(capital) اقتصاد(economy) کاهش(reduction)
religion	اسلام(Islam) آنان(they) خدا(god) حضرت(Hazrat) زنان(women)
sport	فوتبال(football) باشگاه(club) پرسپولیس(Prespolis=a very popular football club) مالی(financial) غرب(west)
Law and politic	زنان(women) دکتر(doctor, Dr.) ماده(bill) یارانه(subsidies) مسکن(housing)
law	ماده(bill) اجرایی(executive) قانونی(legal) جلسه(session, meeting) تصویب(approval)
economy	ارز(currency) مرکزی(central) نرخ(rate) دلار(dollar) فروش(sale)

Table **4** reveals that the utilization of ICA on word-document matrix has led to the detection of syntagmatic word clusters. For example, in one of the extracted clusters, some economically related words namely {“consumption”, “increase”, “capital”, “economy”, “reduction”} are put in the same group. All the aforementioned words could have been used in the same context such as economic related articles.

3.1.3. Extracting features from word-word matrix of Alef dataset

500 most frequent words form the dictionary, and the context words are 2000 most frequent words. Thus, word-word matrix X for this dataset is a 500-by-2000 matrix. The words we want to cluster are represented by vectors in a 2000 dimensional space. 10 components are extracted from this word-word matrix, and each component is a 2000 dimensional vector which represents a cluster. The cosine distance is utilized to determine the closeness of a word to each cluster. Finally, for each cluster, five most representative words are reported in the Table **5**. These words are relevant in a sense that they have been occurred in a similar context.

The utilization of ICA on word-word matrix has led to the extraction of both paradigmatically and syntagmatically related words. For example, collocations such as {Zionist, Regime}, {inter, national}, {principal, law=constitution}, and {Ahmadi, Nejad} embody syntagmatic relations. On the other hand, extracted words such as {affair, case, problem, background} strongly exemplify a paradigmatic synonymous relation.

Table 5: Automatic thesaurus construction using word-word matrix extracted from the Alef corpus	
Automatic category	Words selected by WordICA method
management	شود(become) ریاست(executive) شورای(council) گفتگو (discussion) جلسه(session, meeting, convention)
government	صهیونیستی(Zionist) رژیم(Regime) خارجه(foriegn) وزیر(minister) امور(affair)
political	المللی(national) بین(inter) ایران(Iran) نفت(oil) هسته(core, nuclear)
currency	تومان(Toman=Iranian currency) هزار(thousand) میلیارد(billion) میلیون(million) حدود(limitations, about, nearly, almost)
Law	قانون(rule) اساسی(principal) برای(for, in order to) اجرای(implementation) بین(inter)
Synonyms of affair	این(this) امر(affair) موضوع(case, subject) مسئله (problem) زمینه(background)
Ahmadi Nejad	احمدی(Ahmadi) نژاد(Nejad) آقای(Mr) دکتر(Doctor) جمهور(public)
government	اسلامی(Islamic) جمهوری(republic) شورای(council) ایران(Iran) مجلس(parliament)
law	حالی(a condition) قانون(law, legislation) المللی(national) اساسی(principal) ماده(material, bill=law proposal)
Time	سال(year) گذشته(last) جاری(current) ماه(month) آینده(future, coming, next)

3.1.4. Extracting features from word-document matrix of Tabnak dataset

First of all, a dictionary of all the words in this dataset is created. The words are ordered in terms of frequency so that the most frequent word is ranked the first word. Since a few of the most frequent words do not convey much information, they are not considered in the experiment. The analyzed words are from the list of words ranked from 201 to 700. This will lead to a 500-by-20872 matrix. ICA is applied to this matrix, and 10 independent components are extracted. The values of each component denote the word features. Each one of 500 words is compared with each one of the 10 components. In other words, 5000 comparison is accomplished. Finally, 5 closest words to each component are selected. They can be the most representative 5 words for a cluster. These ten automatically extracted clusters are reported in the Table **6**.

By applying WordICA on a word-document matrix extracted from the Tabnak corpus, clusters are detected in which the words are mostly syntagmatically related. For instance, the words namely {“Turkey”,”Europe”,”Union”,”relationship”,”security”} are strongly related in a syntagmatic manner.

Although most of the extracted words are related, they might seem semantically far from each other from a person’s point of view who is unfamiliar to the Iranian culture. I shed light on the possible relation among the words which belong to the last cluster namely {Sepah, culture, commander, language, Hussein}. After the 1979 Iranian revolution, a revolutionary army is founded which was called Sepah [129]. It was first governed autonomously, but then became a branch of Iran’s military. According to the Iran’s constitution, its objective is to defend the Islamic system of Iran. This unequaled military army is active in diverse areas including culture. Furthermore, this army is greatly inspired by Hussein [130] (grandson of Prophet Muhammad) who is considered to be the most influential figure in Shiite Islam. Imam Hussein has significantly impressed the culture, language, and political view of Iranians and other Islamic nation. That’s how these words become syntagmatically related in a religious and political context.

3.1.1. Extracting word features from word-word matrix of Tabnak dataset

500 most frequent words are analyzed for finding word clusters. The contextual information is collected from 2000 most frequent words. In other words, the word-word X matrix is 500-by-2000. 10 independent components are extracted from this matrix. Each one of these component is an abstract representation of a word cluster. Each one of the 500 analyzed words is compared with these 10 components, and the words which are closest to a component may form a cluster. You can find these 10 word clusters in the Table **7**.

Table 6: Automatic thesaurus construction using word-document matrix extracted from the Tabnak corpus

Detected category	Extracted words
adverbs	خیلی(very) خودش(himself) چیزی(something) مثل(like, similar) درست(right, exact)
politic	ترکیه(Turkey) اروپا(Europe) اتحادیه(union) روابط(relationship) امنیتی (security)
football	استقلال(Esteqlal) پرسپولیس(Prespolis) مهدی(Mahdi, who is possibly a famous football player) رضا(Reza, who is also another football player) پیروزی (Piroozi=another name for Prespolis, which is a famous football team; the second meaning=victory)
judicial	قضایی(judicial) ماده(material, bill=law proposal) قوه(system, flashlight) رسیدگی(considering) ازدواج (marriage)
religion	قرآن(Quran) دین(religion) کتاب(book) انسان(man, human) خدا (God)
politic	سپاه(Sepah= Army of the Guardians of the Islamic Revolution) مصر(Egypt) زبان(language) فلسطین(Palestine) اجلاس (convention)
war	لبنان(Lebanon) ترور(assassination) مقاومت(persistence) امنیتی(security) ارتش (army)
sport	آژانس(agency) لیگ(league) دادگاه(court) فدراسیون(federation) تابناک (Tabnak)
sport	ورزش(sport) فدراسیون(federation) جوانان(young people) مسابقات(competitions) کمیته (committee)
War and politic	سپاه(Sepah= Army of the Guardians of the Islamic Revolution) فرهنگ(culture) فرمانده(commander) زبان(language) حسین (Hussein= name of a person)

In Table **7** the application of WordICA on word-word matrix extracted from the Tabnak corpus has led to the detection of both paradigmatic and syntagmatic word clusters. For instance, the words in the second cluster namely {“is”, “has become”, “has not become”, “has arrived”, “has been”} are all instances of paradigmatic synonyms. There are also other paradigmatically related words such as {“got”, “has gotten”}, {“leader”, “leadership”}, {“in order to”, “so that”}, {“this”, “that”}, and {“Imam”, “Hazrat”}. Both “Hazrat” and “Imam” are glorious Arabic title used to honor a (usually a prophet or a saint) person. On the other hand, several syntagmatically related words have been excavated, like {“Zionist”, “Regime”, “foreign”, “minister”, “affair”}, and {“Islamic”, “Republic”, “council”, “parliament”, “Iran”}.

3.2. Motion data analysis

In this part of the project, the annotations of motion videos collected in February from the online web-based form is analyzed. This form is for annotating motion videos using English, Finnish, Swedish, Farsi, or any other language. Each one of the 124 videos has been transformed to a 602-dimensional vector using the motion data [142]. Besides that, every word, which has been used for annotating a video, has also been converted to a 124-dimensional vector. The videos and the words used for annotating them have been depicted on the same 2-dimensioal plane.

Table 7: Automatic thesaurus construction using word-word matrix extracted from the Tabnak corpus	
Automatic category	Words selected by WordICA method
	(to get)گیرد (relax)قرار (day)روز (Hazrat)حضرت (Imam)امام
Light verbs	(has arrived)رسیده (has not become)نشده (has become)شده (is)است (has been) بوده
Light verbs	(exposed to, case, about)مورد (got)گرفت (to get)گیرد (relax)قرار (has gotten)گرفته
Adjectives for supreme leader of Iran	(leadership)رهبری (Maqam, dignity, rank)مقام (Moazzam, supreme)معظم (leader) رهبر (this)این
Foreign affair	(affair)امور (minister)وزیر (foreign)خارجه (Zionist)صهیونیستی (regime)رژیم
Linking words	(this) این (be able to)تواند (so that, which)اینکه (that)آن (for, in order to)برای
	(security) امنیت (condition)حالی (council)شورای (get)گیرد (relax)قرار
Foreign affair	(be able to) تواند (minister)وزیر (ministry)وزارت (foreign)خارجه (affair)امور
Governmental	(Iran) ایران (parliament)مجلس (council)شورای (Republic)جمهوری (Islamic)اسلامی
Adjectives for supreme leader of Iran	(revolution)انقلاب (leader)رهبر (dignity)مقام (leadership)رهبری (supreme)معظم

3.2.1. Verbs and adjectives distribution

There are 124 videos in the form. People who fill the form are not limited to use a select few words. An interesting research question is to see whether people have different opinions about the meaning of a verb. This has propelled us to visualize the variety of people's choice on verbs and adjectives. In order to do so, we have extracted the frequency of verbs and modifiers from the form. In addition, the motion data such as limb end position and the velocities of the animated characters have been calculated. As a result 602-dimensional data points have been created which cannot be visualized. Therefore, PCA is utilized as a dimensionality reduction method, and only the first two principal components have been considered to convert it to a two-dimensional data which can easily and intuitively be analyzed.

In the following figures, each video is represented by a pie. The position of the pie has been determined by the first two principal components of the high dimensional motion data calculated from the raw motion capture data coming from the visual modality. Besides that, the size of a pie reflects the number of verbs assigned to that video. The color reflects different selection of words. If videos with similar color distributions are put close to each other in a densely packed area in this newly created space, it would indicate that similar videos can be clustered well even in this two-dimensional space. Thus, the features we have selected can be utilized in the future to map the words to actions. In other words, the perfect scenario for automated animation design would be having small clusters with different colors for each cluster.

Figure 7 demonstrates the distribution of the Farsi verbs assigned to motion videos. As it can be seen in the figure, the verb "walking" has occupied most of the space; in other words, there is a large variation in how people use the verb "walking" in Farsi. This indicates that building an optimal motion search based only on verbs is such a tremendous task as the subjectivity of verbs should also be considered.

Furthermore, people are unanimous about the verbs. Limping is less used than the first two verbs, and it occupies the lower left corner of the figure; running has occupied the right and middle part of the space. Besides that, if we forget about the walking, running and limping can be clustered very easily. In other words, people are unanimous whether an action is running or limping.

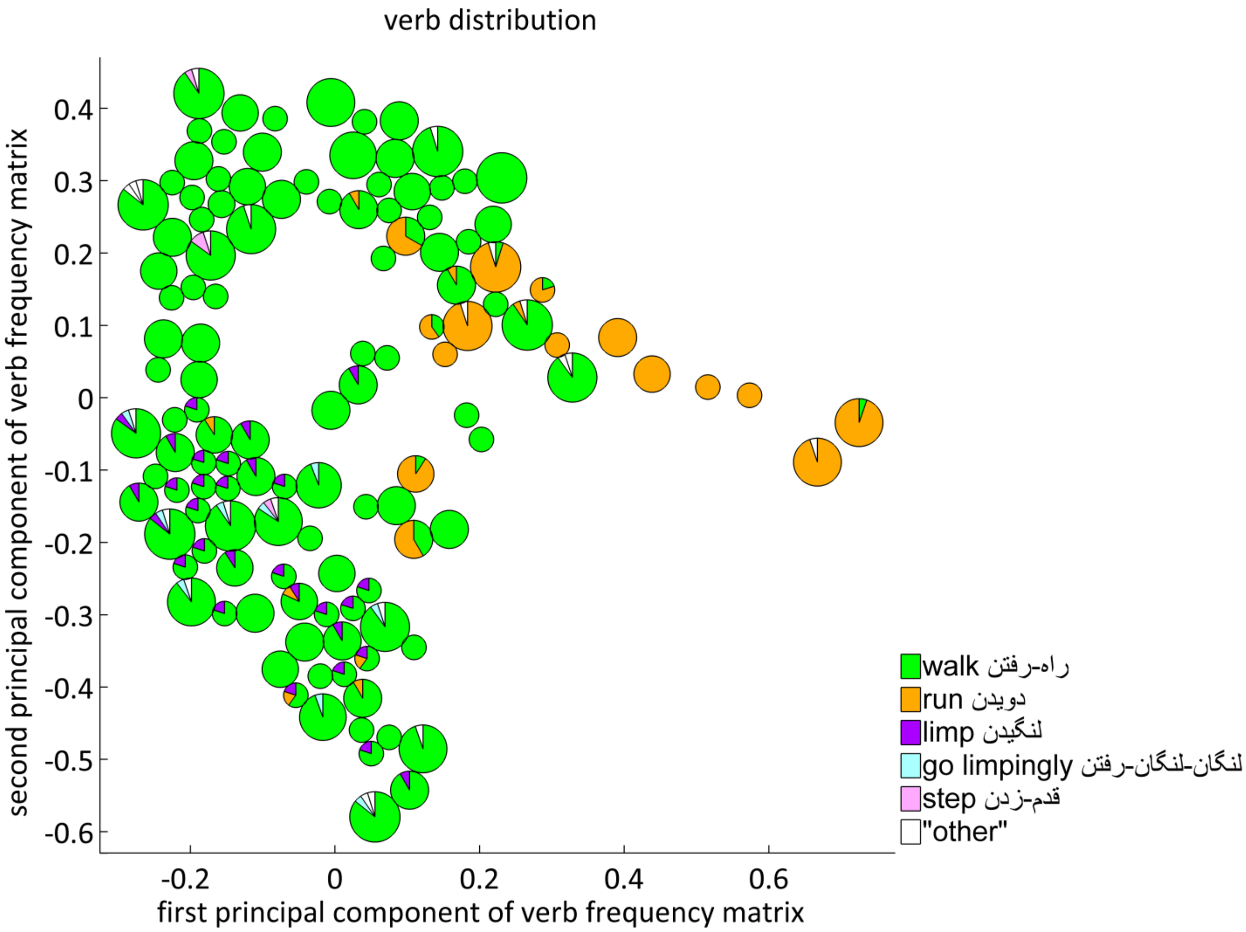


Figure 7: Distribution of the 5 most frequent verbs. Each pie corresponds to one video; its position is determined by the first and second principal component of the visual data; its size is proportional to the number of verbs assigned to that video.

Figure 8 shows that people are less unanimous in their selection of a modifier than a verb. In other words, an action can be described by many modifiers. The adverb "slowly" appears in most part of the space while it is concentrated in the left upper corner of the space. On the other hand, limping is in the left lower corner, and synonyms of fast lie in the right part of the space. Some adverbs have been qualified by other adverbs such as "very" and "a little". These kinds of modifiers do not change the meaning. There are also some synonyms such "تند", and "سریع". This has propelled me to group adverbs into some clusters so that only one of them will be the representative of the whole cluster.

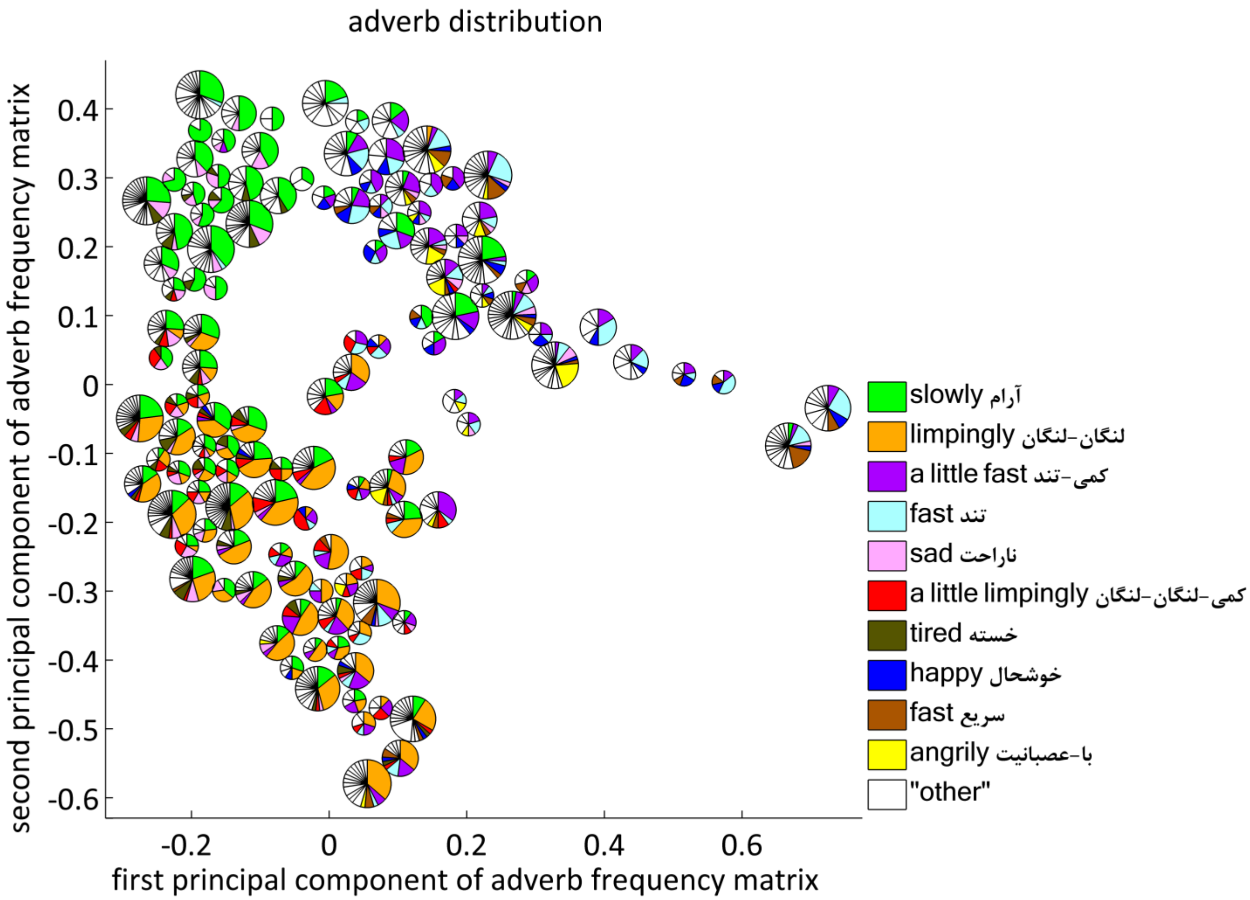


Figure 8: distribution of the 10 most common modifiers. The position of a pie is determined by the two principal components. The size of a pie reflects the number of answers given to its corresponding video, and the position denotes the style.

Figure 9 demonstrates the distribution of verbs combined with adjectives. Among these 10 most frequent combinations, the ones that are synonym of “limping” have occupied the bottom left corner of the space, while the synonyms of “walking” are mostly in upper left corner of the space. In addition, people are less unanimous about whether a motion is running or walking fast.

3.2.2. Grouping adverbs

There are many adverbs used to describe a motion, so by manually grouping them into clusters, we can have a better understanding of its visualization. The process of segmentation is a heuristic approach. The adverbs list has been observed at first. Some of the adverbs can be said in different ways. For example, limping can be expressed in four different ways:

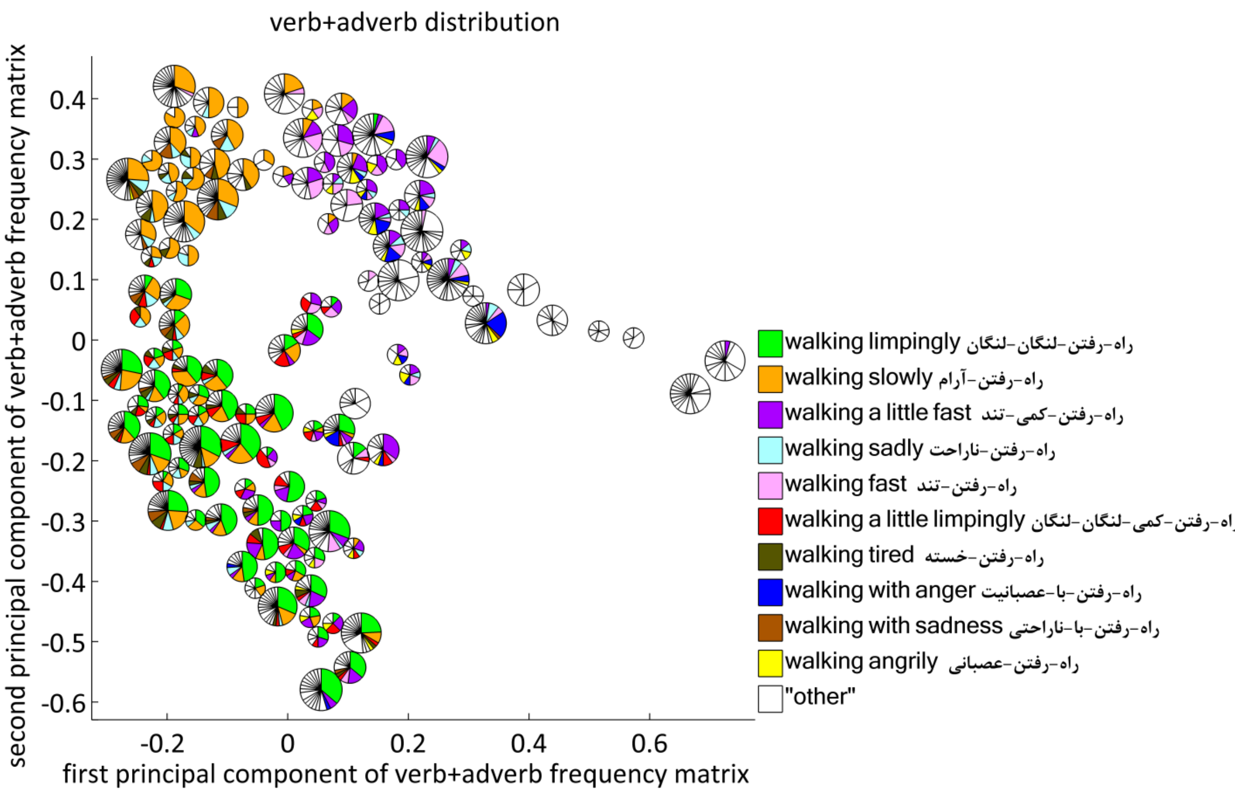


Figure 9: distribution of the 13 most common frequent combinations of verb and its modifier.

1. لنگ_لنگان
2. لنگ_لنگ_زنان
3. لنگان_لنگان
4. لنگان

In addition, two other used words are synonyms of limping:

1. کشان_کشان
2. با_شلیدن

In addition to different ways of expression and synonyms, morphemes have also been considered. Derivational morphemes can be added to adjectives for converting it into an adverb. For instance, the adjective “خوشحال” can be converted to an adverb by adding one morpheme to its beginning and one morpheme to its end.

Table 8: Composition of the adverb “happily” in Farsi			
Word	morphemes	Meaning	Part of speech
با خوشحالی	با+خوشحال+ی	With happiness/happily	Adverb
خوشحال	خوشحال	happy	adjective

Annotators have utilized both adjectives and adverbs to describe a motion; some of the adverbs can be easily reduced to adjectives if we remove their derivational morphemes and keep only the roots. However, we should notice that the first morpheme can be written immediately next to the root or with one space from it. This space can be a zero-width-non-joiner or a full space.

Furthermore, there are modifiers such as “خيلي (=very)” and “كمي (=a little)” in Farsi whose role is to qualify other adverbs; if we remove these kinds of modifiers the meaning of a sentence is not hurt.

All things considered, six groups of adverbs have been created manually by considering synonyms and different ways of expression, while removing meaningless morphemes and modifiers. These six groups correspond to these six adverbs:

1. آرام=slow
2. عصبانی=angry
3. بیحال=languid
4. ناراحت=sad
5. سریع=fast
6. طبیعی=normal

Other groups could have also been constructed, but because their corresponding adverbs were very rare, they have not been segmented into clusters.

Figure 10 represents the distribution of adverbs after combining the six most frequent adverbs into groups. Although the majority of the annotators were more unanimous about the adverb “fast”, a few motions have been labeled using the opposite adverb “fast” and “slow”. These cases might be caused by expressing the same motion using different verbs. Those annotators who have described these motions as walking might have utilized the adverb “fast” to distinguish it from the regular walking. On the other hand, those people who have described the same action as running might have utilized the adverb “slow” to differentiate it from regular running. In addition, the term “fast” appears on the right side of the line x=-0.1, while the term “limping” appears below the line y=0.1. It is worth mentioning that such information is obtained from a two-dimensional representation of a high dimensional data. Furthermore, sadness has been used with slowness, while happiness has come mostly with fastness. Anger is also associated with fastness.

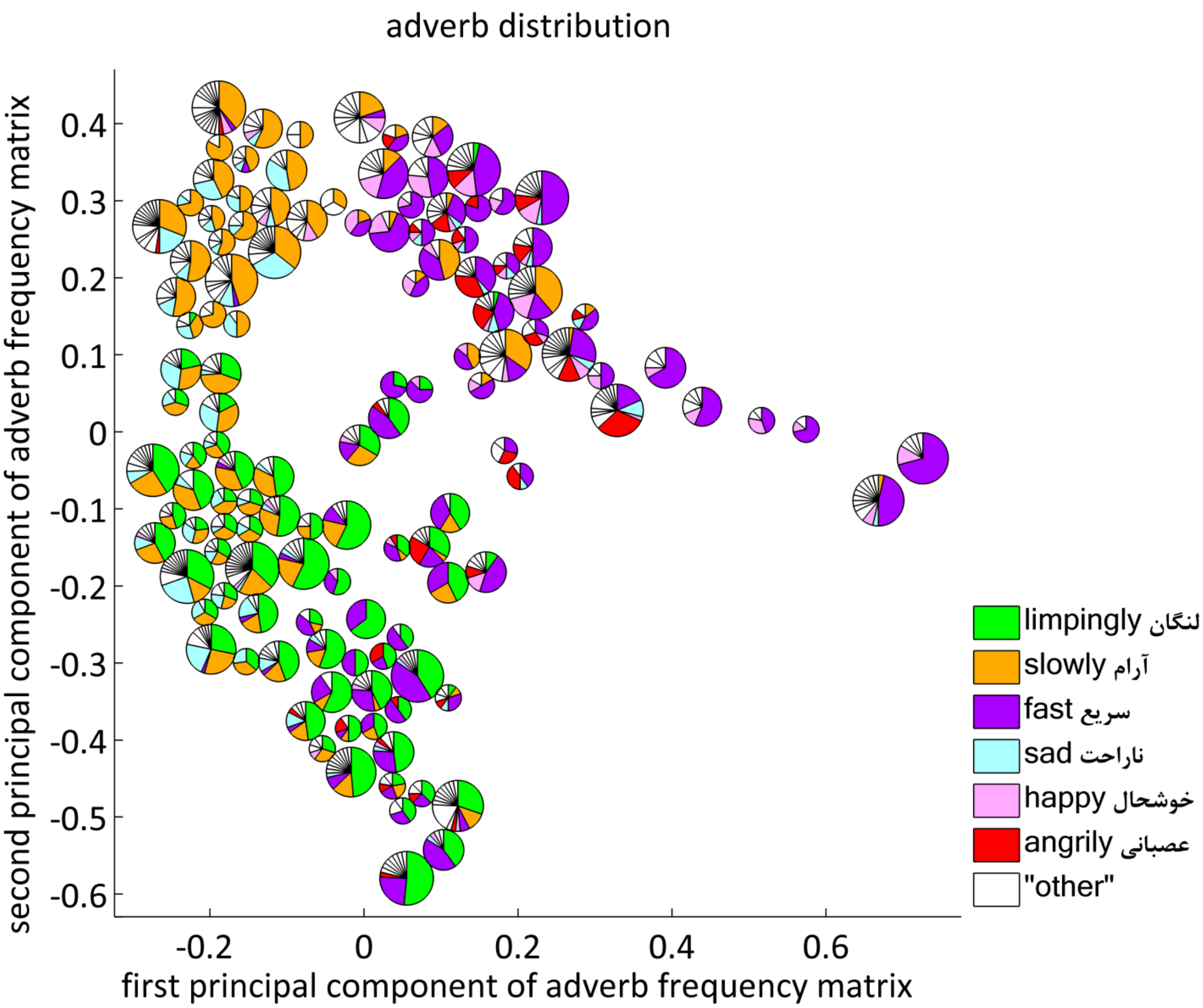


Figure 10: distribution of the 6 most common adverbs. Each adverb represents a group of adverbs which have been segmented by removing meaningless morphemes and modifiers; also, different ways of expression and synonyms have been considered.

Figure 11 shows the distribution of the verb-adverb combination. The majority of the annotators are unanimous about “running fast”. Besides that, “walking sad” has appeared mostly with “walking slowly”, while “walking angrily” has come mostly with “walking fast”, which reveals how these motions and emotions are related to each other.

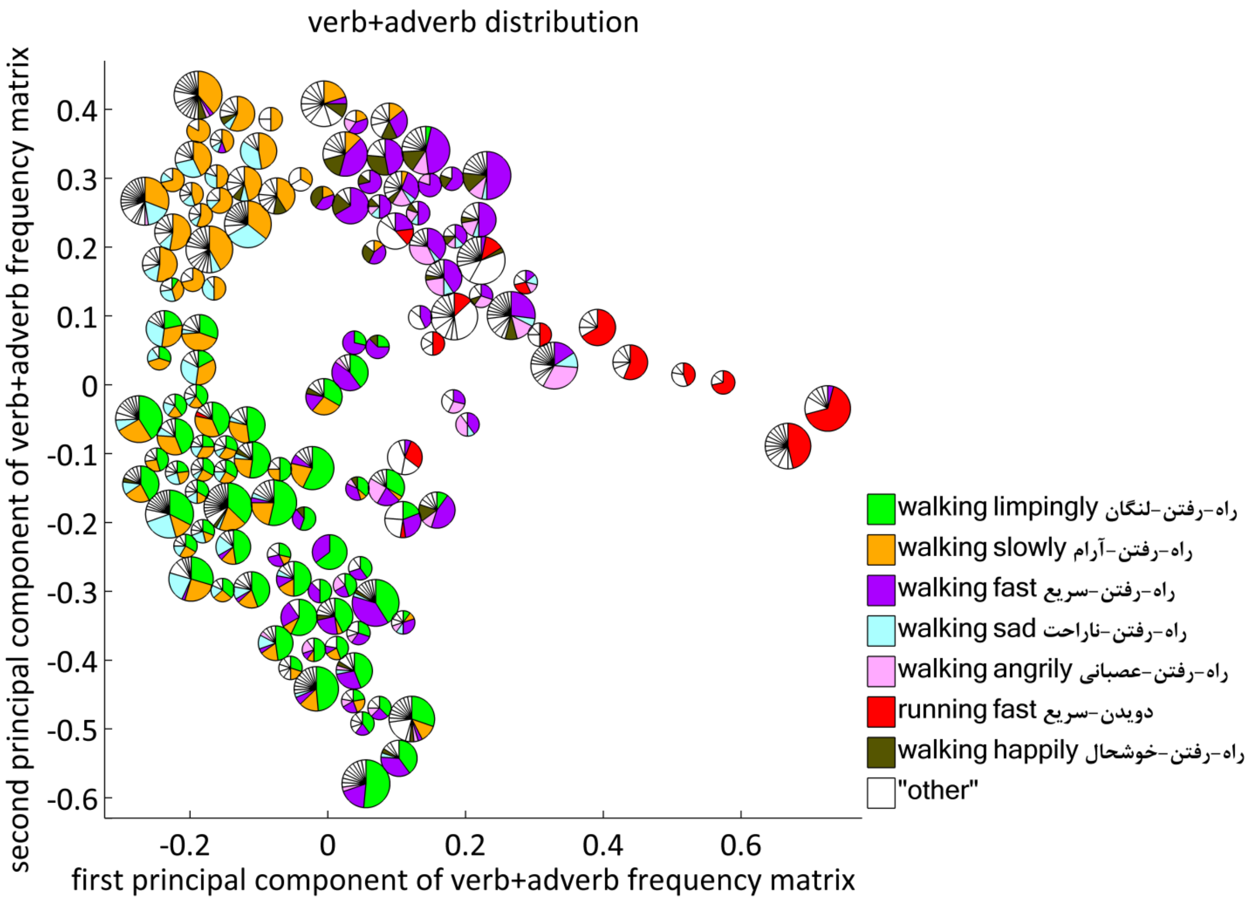


Figure 11: distribution of the 7 most verb-adverbs combination.

Figure 12 shows the distribution of the 10 most common English annotated verbs. By comparing this figure, the previous figures demonstrated in this chapter, and the figure 5 of chapter 5 from [142], we can see that synonyms of {'walking', 'راه رفتن', 'kävelee'} has occupied the upper left corner of the 2-dimensional space; while synonyms of {'limping', 'لنگیدن', 'ontuu'} lie in the lower left corner, and the synonyms of {'running', 'دویدن', 'juoksee'} are in the right part of the figures. This similar distributional pattern prompted us to use the video index as the context for transforming verbs and verb-adverbs into vectors. Now that we have a numerical representation for verbs and verb-adverbs, we can find synonyms and translations. Furthermore, synonyms have emerged very close to each other; as a result, by exploiting the motion data, coming from the visual modality, to normalize the verb and verb-adverb vectors, one can expect to extract more appropriate synonyms and translations.

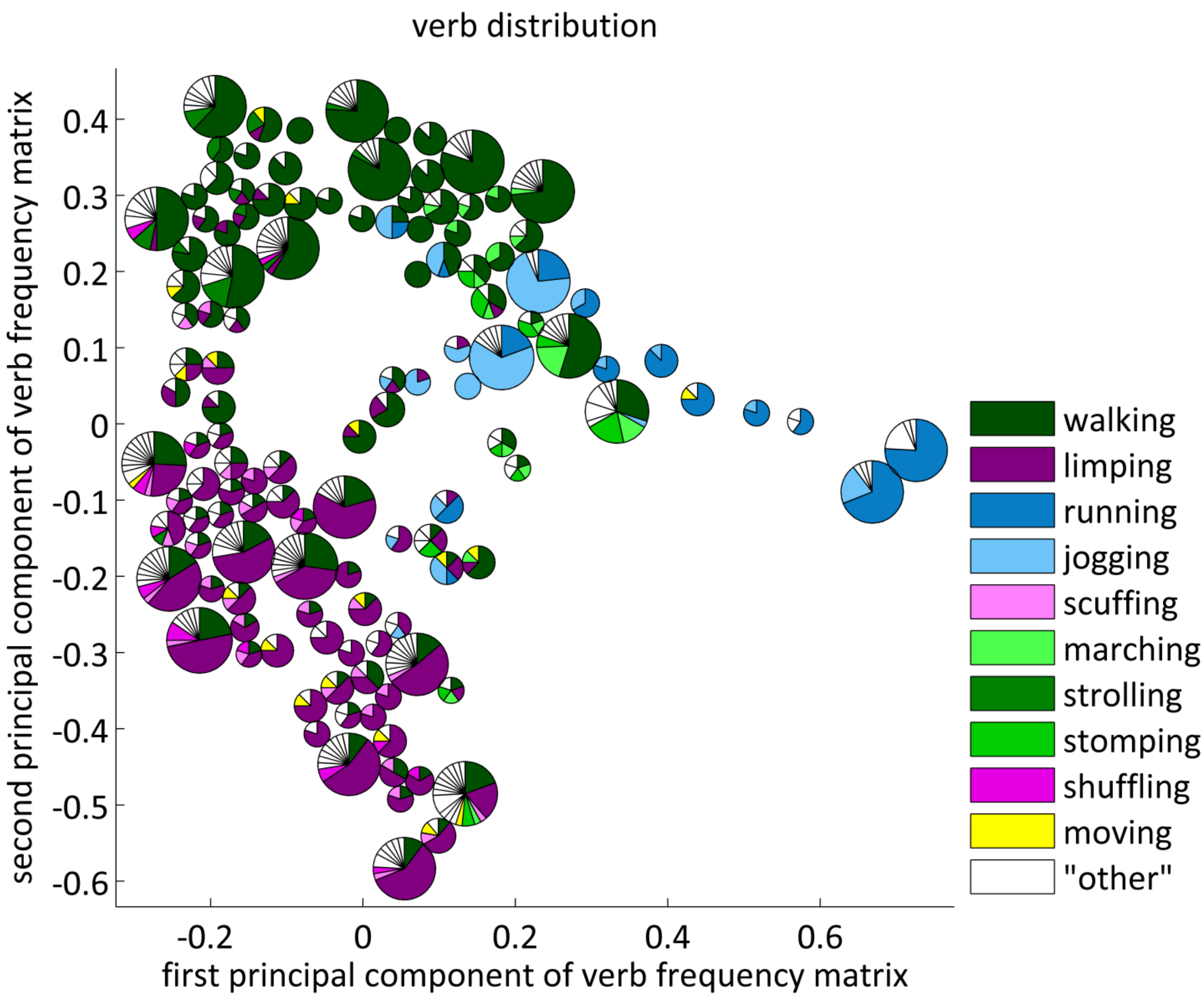


Figure 12: distribution of the ten most common English annotated words

3.3. Synonym results

In addition to symbol grounding using motion data, I have also grounded the verbs and verb-adverbs using a random matrix whose size is exactly the same as the motion data matrix; in this case, the components of the vector defining a video does not have anything to do with the video itself. In fact, the components are nothing but random numbers. Therefore, two relevant videos might be farther than two irrelevant videos because their corresponding vectors have been generated randomly. The purpose of randomized-grounding is to compare it with the indexical-grounding and pattern-grounding. In the following tables, letter '-s' represent the indexical-grounding case; 's' denotes the pattern-grounding circumstance; and, 'rs' displays the randomized-grounding condition.

I have counted the number of good synonyms in each case; by good synonyms, I mean satisfactory apt synonyms; for instance, wandering is a good synonym of walking, while limping is not. On the other hand, if we consider these verbs in a general situation where we have other verbs such as 'swimming', 'eating', 'speaking', 'thinking' or any verbs from other domains excluding the motion realm, we would perceive 'limping' as a synonym of 'walking' or even 'strolling' because they describe certain kinds of movement. However, we have narrowed down our verb selection only to the motion field and excluded other domains.

In the tables of the following sections, you can see the most common verbs and their automatically detected synonyms. I have included only the 3 most frequent words; if you are interested in observing the recognized synonyms of less frequent verbs, you can find them in Appendix D.

3.3.1. English Synonym Results

Table 9: Extracted synonyms of the English motion verbs

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	Number of good synonyms
walking	392	-s	Walking	walk	Strolling	Ambling	wandering	5
		s	walk	Walking	Strolling	amble	wandering	5
		rs	Walking	Limping	walk	limping	scuffing	2
limping	202	-s	Limping	stagger	leaping	climbing	hobbling	4
		s	Limping	scuffing	stagger	limb	leaping	5
		rs	Limping	scuffing	walking	walk	moving	3
Limping	110	-s	limping	leaping	stagger	hobbling	scuffing	5
		s	limping	scuffing	moving	leaping	shuffling	5
		rs	limping	scuffing	walking	walk	Walking	2

First of all, I have to say that we are interested in analyzing the natural language as it emerges; thus, the annotations have not been preprocessed. Hence, small-letter words are distinguished from capital-letter ones. The most frequent verb is 'walking'. In fact, it has been used 392 times for describing a motion video. It seems that all the five closest verbs of 'walking' detected by both indexical-grounding and pattern-grounding approaches are its appropriate synonyms, while in the randomized-grounding approach, only two verbs are its applicable synonyms. At first, it may sound paradoxical that 2 apt verbs have been identified; one may ask himself why both the closest and the third closest word to 'walking' are still its very good synonyms in the randomized-grounding. Well, I have to say at this point, that in the randomized-grounding approach, we are still utilizing the correct video indexes as a word context. As a matter of fact, it is only the normalization process which is implemented randomly. In other words, every component of a word vector stores the original frequency of that word in a video, and the index of each component is precisely mapped to the index of its corresponding video. For instance, if 'walking' has been used to describe the 12^{th} video 34 times, the 12^{th} component of the 'walking' vector is assigned 34 initially. It is only after the normalization step that something is added to 34. This normalization in the indexical-grounding step is based on a meaningful concept, while in the randomized-grounding approach it is baseless. Though, this normalization is implemented softly. Hence, the fact that the closest word to 'walking' is still its relevant synonym is not a contradiction or a big surprise.

Besides considering the synonyms of a verb, the synonyms of the combination of a verb and an adverb are also analyzed. When counting the number of good synonyms for a verb_adverb, the meaning of both the verb and the adverb of the extracted verb_adverb should overlap with meaning of current verb_adverb so that the extracted verb_adverb is identified as a good synonym. For example, in 'ambling_liesurely', the meaning of ambling overlaps with walking, and also the meaning of leisurely overlaps with the meaning of slowly; on the other hand, the meaning of 'joyfully' is very far from 'slowly'. That's why, 'walking_joyfully' cannot be a good synonym of 'walking_slowly'.

Table 10: Extracted synonyms of the English motion verb-modeifiers

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms
walking_slowly	57	-s	walking_thoughtfully	walking_unpurposefully	walk_slowly	Walking_Joyfully	Ambling_Leisurely	2
		s	walking_very_slowly	walking_sadly	walking_carefully	Strolling_Slowly	walking_thoughtfully	2
		rs	Walking_Slowly	walking_sadly	walking_carefully	walking_very_slowly	Limping_Slowly	2
limping_slowly	28	-s	Hobbling_Slowly	limping_sadly	walking_asymmetrically	shuffling_wounded	scuffing_painstakingly	1
		s	limping_sadly	limping_painfully	scuffing_painstakingly	scuffing_slowly	Hobbling_Slowly	2
		rs	limping_painfully	walking_carefully	Limping_Slowly	walking_asymmetrically	scuffing_painstakingly	1
limping_painfully	21	-s	limping_slowly	Limping_Slowly	walk_injured	limping_sadly	Walking_Watchfully	0
		s	limping_slowly	Limping_Slowly	limping_sadly	limbing_very_slowly	Limping_Painfully	1
		rs	limping_slowly	Limping_Slowly	walking_carefully	walking_very_slowly	Limping_Normally	0

The most common verb_adverb combination among English annotations has been 'walking_slowly'. The adverb 'slowly' has emerged 57 times with the verb 'walking'. It does not mean that 57 distinct videos have been annotated by 'walking_slowly'. Instead, it illustrates that an annotator has used this verb_adverb to describe a certain set of videos; then another person might have annotated a different set of videos by 'walking_slowly'; the total frequency of 'walking_slowly' has been 57. In addition, since the frequency of 'walking' is much bigger than the frequency of 'walking_slowly', the vector which represents 'walking_slowly' is sparser than the one representing 'walking'. As a result, the number of good detected synonyms for verb_adverbs is less than good synonyms for verbs.

It is worth noticing that the pattern-grounding approach has managed to detect the best synonym for 'walking_slowly' as closest verb_adverb. In other words, if we just focus on the closest extracted verb_adverb, we can see more conspicuously that the pattern-grounding approach beats the other two approaches. The second most frequent verb-adverb is 'limping-slowly'. Although, the pattern-grounding approach did not manage to detect its best synonym as the closest one, it has found more good synonyms than the other two approaches. In the indexical-grounding approach, the only factor which determines the closeness of two words is the co-occurance in the same context. For example, 'Hobbling-Slowly' has co-ocurred frequently with 'limping-slowly'; that's why, they have been categorized as synonyms. On ther other hand, in pattern-grounding approach, the proximity of the context is also analyzed. As a result, 'Hobbling-slowly' has moved down to the fifth closest, while 'scuffing-slowly' which has not been identified as a synonym previously, is correctly categorized as a synonym with the help of pattern-grounding.

3.3.2. Farsi synonym result

For Farsi annotations, only the synonyms of verb-adverbs have been extracted because the Farsi verbs used for annotating the motion videos have much less diversity than the English or Finnish motion verbs.

Table 11: Extracted synonyms of the Farsi motion verb-modifiers

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms
راه_رفتن_آرام	389	-s	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_خسته	راه_رفتن_غمگین	راه_رفتن_بی_حوصله	1
		s	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_خسته	راه_رفتن_غمگین	راه_رفتن_خیلی_آرام	2
		rs	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_خسته	قدم_زدن_اهسته	راه_رفتن_لنگ_لنگان	2
راه_رفتن_لنگان_لنگان	346	-s	راه_رفتن_لنگ_لنگان	راه_رفتن_لنگان	لنگیدن_خسته	لنگیدن_آسیب_دیده	لنگان_لنگان_راه رفتن_خیلی_لنگیدن	3
		s	راه_رفتن_لنگ_لنگان	قدم_زدن_بالنگیدن	راه_رفتن_کمی_لنگان_لنگان	قدم_زدن_با_لنگیدن	لنگیدن_آسیب_دیده	4
		rs	راه_رفتن_لنگ_لنگان	راه_رفتن_کمی_تند	راه_رفتن_کمی_لنگان_لنگان	راه_رفتن_خسته	لنگیدن_خسته	2
راه_رفتن_کمی_تند	135	-s	راه_رفتن_تند	راه_رفتن_خوشحال	راه_رفتن_سریع	راه_رفتن_عصبانی	راه_رفتن_باعجله	3
		s	راه_رفتن_تند	راه_رفتن_خوشحال	راه_رفتن_سریع	راه_رفتن_باعجله	راه_رفتن_عصبانی	3
		rs	راه_رفتن_تند	راه_رفتن_لنگ_لنگان	راه_رفتن_لنگان_لنگان	راه_رفتن_سریع	راه_رفتن_کمی_لنگان_لنگان	2

When pattern-grounding approach is applied to find the closest verb-adverb vectors, more appropriate synonyms are found for Farsi verb-adverbs. Furthermore, some of the apt synonyms get a better rank in terms of its proximity to the current verb-adverb. For instance, 'راه_رفتن_باعجله' is ranked the 5th closest vector to 'راه_رفتن_کمی_تند' by indexical-grounding approach, while it is ranked the 4th closest vector when pattern-grounding is exploited.

3.3.3. Finnish synonym result

The synonyms of Finnish verbs and verb_adverbs have been extracted from the annotations by considering the video indexes in which they have been used as the context. Next, they have been grounded using motion data. Finally, the vectors corresponding the Finnish verbs and verb-adverbs have also been normalized using random data. The following tables include the synonyms of the three most frequent Finnish verb and verb-adverbs.

In Table 12, we can see that both indexical-grounding and pattern-grounding approach has a good performance in finding relevant synonyms. For example, all five extracted synonyms of the most frequent verb, 'kävelee', are its correct synonyms. On the other hand, when randomized-grounding approach is applied, only one apt synonym is found.

Table 12: Extracted synonyms of the Finnish motion verbs

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms
kävelee	618	-s	kävellä	käppäilee	käveleminen	Kävelee	löntystelee	5
		s	kävellä	käppäilee	Kävelee	löntystelee	maleksii	5
		rs	kävellä	nilkuttaa	ontua	ontuu	linkuttaa	1
ontuu	229	-s	nilkuttaa	ontua	linkuttaa	laahustaa	raahustaa	3
		s	ontua	nilkuttaa	linkuttaa	raahustaa	laahustaa	3
		rs	nilkuttaa	ontua	linkuttaa	laahustaa	kävelee	3
nilkuttaa	164	-s	ontuu	linkuttaa	ontua	linkkaa	liikkuu	4
		s	ontuu	linkuttaa	ontua	raahustaa	laahaa_jalkaa	4
		rs	ontuu	ontua	linkuttaa	kävelee	kävellä	3

Table 13: Extracted synonyms of the Finnish motion verb-modifiers

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms
kävelee_hitaasti	64	-s	kävellä_hitaasti	kävelee_surullisena	kävelee_epävarmasti	kävelee_hyvin_hitaasti	käveleskelee_mietteliäästi	2
		s	kävellä_hidas	kävelee_hyvin_hitaasti	kävellä_hitaasti	kävelee_surullisena	maleksii_hitaasti	4
		rs	kävellä_hitaasti	kävelee_surullisena	kävelee_mietteliäästi	kävelee_varovasti	kävelee_rennosti	1
kävelee_reippaasti	52	-s	kävellä_rivakasti	kävelee_määrätietoisesti	kävelee_päättäväinen	kävelee_kävelee_normaalisti	Kävelee_Reippaasti	4
		s	kävelee_määrätietoisesti	kävelee_päättäväinen	kävellä_rivakasti	kävellä_reippaasti	kävellä_rento	4
		rs	kävelee_määrätietoisesti	kävelee_tavallisesti	kävelee_päättäväinen	kävellä_reipas	kävellä_rivakasti	4
ontuu_hitaasti	47	-s	nilkuttaa_hitaasti	ontuu_varovasti	ontuu_vaivaloisesti	nilkuttaa_kivulloisesti	linkuttaa_vaivainen	1
		s	nilkuttaa_hitaasti	ontuu_varovasti	ontuu_hyvin_kivuliaasti	ontuu_vaivaloisesti	ontuu_takavetoisesti	1
		rs	nilkuttaa_hitaasti	ontuu_varovasti	kävelee_vaivalloisesti	ontuu_vaivaloisesti	ontuu_pahasti	1

Four relevant synonyms for the most frequent Finnish verb-adverb have been extracted using the pattern-grounding approach; it is only the 4th extracted synonym which is irrelevant. Pattern-grounding also performs more accurately when finding the synonyms of 'kävelee_reippaasti'; in fact, this approach yields in four apt synonyms with only the 5th extracted synonym being irrelevant, while the irrelevant synonym is ranked the 4th closest synonym using indexical-grounding and the 2nd by randomized-grounding.

3.3.1. All synonyms results

In the following table, you can observe the number of relevant synonyms extracted using all three different approaches. For every verb and verb-adverb, five potential synonyms have been found. The potential synonyms are ranked according to their proximity to the studied verb or verb-adverb. If the meaning of a potential synonym is significantly close to the meaning of the studied verb, it is considered as a relevant synonym of that verb. The potential synonym of a verb-adverb is relevant when the meaning of both the verb and adverb are closely related. With these constraints in mind, the potential synonyms of the 10 most frequent verbs and verb-adverbs have been collated. The following table demonstrates just the numbers; the actual synonyms of the 3 most frequent verbs and verb-adverbs have been represented in previous sections. An interested reader can also visit Appendix D for observing all extracted synonyms.

Table 14: All synoym results

	Number of relevant synonyms(out of 50)		
	indexical grounding	pattern grounding	Randomized-grounding
Farsi verb-adverbs	25	28	18
English verbs	44	46	25
English verb-adverbs	17	19	12
Finnish verbs	40	40	21
Finnish verb-adverbs	28	34	22

Table 14 demonstrates that pattern-grounding slightly improves automatic synonym detection of verbs and verb-adverbs from Farsi, English and Finnish annotations. On the other hand, when grounding is based on random data, the performance of finding relevant synonyms deteriorates. Hence, it is safe to say that pattern-grounding is a meaningful and reliable process which can be utilized to enhance the synonym detection task.

3.4. Translation results

Translation of both verbs and verb-adverbs among English, Finnish, and Farsi language has been implemented using indexical-grounding, pattern-grounding, and randomized-grounding approach, and the number of good translations have been counted in each case. In the following sections, the translations of the three most frequent verbs and verb-adverbs have been covered. If interested in observing all translation result, you are recommended to check Appendix E.

3.4.1. Translation of English annotations to Finnish

Both the indexical-grounding and pattern-grounding approach has resulted in five good translations for 'walking'. On the other hand, the result of randomized-grounding is not so good. The indexical-grounding approach is qualified slightly better for translating 'limping'. The reason for this attainment is that the fifth closest translation is 'liikkuu' which is a general verb that can describe any motion. The fifth closest translation extracted by pattern-grounding is 'raahustaa' which should be translated to ‘shamble’ and not ‘limping’. Yet, using pattern-grounding approach, the less frequent English verbs are translated more accurately to Finnish.

Table 15: Translation of the English to Finnish motion verbs

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	#good translation
walking	392	-s	kävelee	kävellä	Kävelee	käveleminen	käppäilee	5
		s	kävelee	kävellä	Kävelee	käppäilee	astelee	5
		rs	kävelee	kävellä	ontuu	nilkuttaa	ontua	2
limping	202	-s	ontuu	nilkuttaa	linkuttaa	ontua	liikkuu	5
		s	nilkuttaa	ontuu	linkuttaa	ontua	raahustaa	4
		rs	ontuu	nilkuttaa	linkuttaa	ontua	kävelee	4
Limpin g	110	-s	ontuu	nilkuttaa	linkuttaa	ontua	liikkuu	5
		s	ontuu	nilkuttaa	ontua	linkuttaa	raahustaa	4
		rs	ontuu	nilkuttaa	ontua	linkuttaa	kävelee	4

Table 16: Translation of the English to Finnish motion verb-modifers

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
walking_slowly	57	-s	kävelee_hita asti	kävellä_hita asti	löntystelee_r ennosti	käveleskelee _mietteliääst i	kävelee_renn osti	4
		s	kävelee_hita asti	kävellä_hida s	kävellä_hita asti	kävelee_rau hallisesti	kävelee_hyvi n_hitaasti	5
		rs	kävelee_hita asti	kävellä_hita asti	kävelee_ren nosti	kävelee_rau hallisesti	ontuu_hitaas ti	4
limping_slowl y	28	-s	ontuu_hitaas ti	nilkuttaa_hit aasti	ontuu_raska asti	ontuu_alakul oisesti	nilkuttaa_kiv ulloisesti	2
		s	ontuu_hitaas ti	nilkuttaa_hit aasti	ontuu_varov asti	ontuu_vaival loisesti	laahustaa_on tuen	3
		rs	ontuu_hitaas ti	ontuu_varov asti	ontuu_vaival loisesti	nilkuttaa_hit aasti	ontua_hitaas ti	3
limping_painf ully	21	-s	nilkuttaa_kiv ulloisesti	laahustaa_va ivalloisesti	ontuu_hitaas ti	ontuu_väsyn eesti	raahustaa_tu skaisesti	3
		s	nilkuttaa_hit aasti	ontuu_hitaas ti	ontuu_vaival loisesti	ontuu_hyvin _kivuliaasti	ontua_surulli nen	2
		rs	ontuu_hitaas ti	ontuu_vaival loisesti	kävelee_ram miten	ontuu_pahas ti	ontuu_hyvin _kivuliaasti	2

Table 16 shows that using the pattern-grounding approach, the translation of English verb-adverb to Finnish is enhanced. For example, all five Finnish verb-adverbs are appropriate translations for ‘walking_slowly’. Although the correct translations for ‘limping_painfully’ are not highly ranked, using pattern-grounding the correct tranlations for other verb-adverbs are ranked more appropriately, and more correct translations are found. You can find the translation of other verb-adverbs in the Appendix E.

3.4.2. Translation of Finnish annotations to English

Table 17: Translation of the Finnish to English motion verbs

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
kävelee	618	-s	walking	Walking	walk	striding	Ambling	5
		s	walking	Walking	walk	Strolling	amble	5
		rs	walking	Walking	Limping	walk	limping	3
ontuu	229	-s	limping	Limping	scuffing	stagger	leaping	3
		s	limping	Limping	scuffing	shuffling	limb	4
		rs	limping	Limping	scuffing	walking	moving	3
nilkuttaa	164	-s	limping	Limping	stagger	scuffing	leaping	3
		s	limping	Limping	limbing	scuffing	limb	4
		rs	Limping	limping	scuffing	walking	moving	3

In Table 17, Pattern-grounding obviously boosts the translation performance. The number of good translations is counted manually and after human interpretation; in addition, typos have been disregarded. For instance, using pattern-grounding, the 5th closest translation is ‘limb’; it is most likely that an annotator has meant ‘limp’ by typing ‘limb’ because none of the motion video has anything to do with the word ‘limb’; furthermore, the use of ‘limb’ as a verb in ‘limbing’ is another evidence of this logical reasoning.

In Table 18, we can see that pattern-grounding leads to more apt translations for the most frequent Finnish verb-adverb. Although randomized-grounding results into one more applicable translation for ‘kävelee_reippaasti’, its best translation is ‘walking_briskly’ which is ranked the 3rd by randomized-grounding, while it is ranked the 2nd by pattern-grounding. In addition, pattern-grounding approach slightly enhances the translation of less frequent verb-adverbs.

3.4.3. Translation of English annotations to Farsi

Although both indexical-grounding and pattern-grounding approaches has recognized 5 good translations for the English verb ‘walking’, the translations which have been discovered using pattern-grounding are more appropriate as the 4th translation of the indexical-grounding, which is ‘تند_راه_رفتن’, has an adverb which constrain the meaning of the verb. In other words, ‘تند_راه_رفتن’ is made of a verb ‘راه_رفتن’=’walk’ and an adverb ‘تند’=’fast’. On the other hand, all the transations of ‘walking’ found by pattern-grounding are Farsi verbs. The Farsi translation of ‘limping’ can be expressed in various ways. For instance it can be translated into a single Farsi verb such as ‘لنگیدن’ or to a combination of Farsi verb-adverb ‘لنگان_لنگان_راه_رفتن’; in the latter case the adverb has come before the verb. There can also be an appropriate translation where the adverb comes after the verb such as in ‘راه_رفتن_لنگان_لنگان’.

Table 18: Translation of the Finnish to English motion verb-modifers

Verb-adverb	Frequency	Mode	1st extracted verb-adv	2nd	3rd	4th	5th	# good translation
kävelee_hitaasti	64	-s	walking_slowly	walking_thoughtfully	walking_unpurposefully	walking_very_slowly	walking_waiting	3
		s	walking_slowly	walking_very_slowly	walking_thoughtfully	walking_idly	walking_waiting	4
		rs	walking_slowly	walking_very_slowly	walking_carefully	Walking_Slowly	limping_painfully	3
kävelee_reippaasti	52	-s	Walking_Purposefully	walking_erect	walking_energetically	walking_fast	walking_confidently	2
		s	walking_steadily	walking_briskly	Walking_Purposefully	walking_erect	walking_actively	3
		rs	walking_confidently	walking_fast	walking_briskly	walking_steadily	walking_energetically	4
ontuu_hitaasti	47	-s	limping_slowly	limping_sadly	Hobbling_Slowly	shuffling_wounded	limping_painfully	2
		s	limping_slowly	limping_painfully	limping_sadly	scuffing_slowly	scuffing_painstakingly	1
		rs	limping_slowly	limping_painfully	Limping_Slowly	scuffing_slowly	walking_slowly	2

Table 19: Translation of the English to Farsi motion verbs

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
walking	392	-s	راه_رفتن	راهرفتن	قدم_زدن	تند_راه_رفتن	راه_می_رود	5
		s	راه_رفتن	راهرفتن	قدم_زدن	پیاده_روی	گام_برداشتن	5
		rs	راه_رفتن	قدم_زدن	گام_برداشتن	راهرفتن	لنگیدن	4
limping	202	-s	لنگیدن	لنگان_لنگان_رفتن	لنگان_لنگان_راهرفتن	لنگان_لنگان_راه_رفتن	می_لنگد	5
		s	لنگیدن	لنگان_لنگان_راهرفتن	لنگان_لنگان_رفتن	قدم_برداشتن	لنگان_لنگان_راه_رفتن	4
		rs	لنگیدن	لنگان_لنگان_راهرفتن	راه_رفتن	قدم_زدن	گام_برداشتن	2
Limping	110	-s	لنگیدن	لنگان_لنگان_رفتن	لنگان_لنگان_راهرفتن	می_لنگد	لنگان_لنگان_راه_رفتن	5
		s	لنگان_لنگان_راهرفتن	لنگیدن	لنگان_لنگان_رفتن	لنگان_لنگان_راه_رفتن	راه_رفتن	5
		rs	لنگیدن	راه_رفتن	لنگان_لنگان_راهرفتن	قدم_زدن	گام_برداشتن	3

It is worth mentioning that the translations have been found by considering only the verb inputs which annotators have provided. It is also possible that in the verb box of the motion form an annotator has inputted an expression which is made of a verb-adverb combination. 'قدم_برداشتن'= {'to step, to walk'} is not considered as an appropriate translation for 'limping' as the correct adverb constraing the meaning of 'walking' to 'walking limpingly' or to 'limping' is not known. The adverb might have been 'لنگان'='limpingly', but because we are selecting the translation of 'limping' from only Farsi verbs, 'قدم_برداشتن' shall not be considered as a good translation of 'limping'. As a result, the indexical-grounding approach has a little better performance in translating the English verb 'limping' to Farsi verbs. However, pattern-grounding prevails in terms of finding more appropriate translations for less frequent English verbs.

Table 20: Translation of the English to Farsi motion verb-modifiers

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
walking_slowly	57	-s	راه_رفتن_آرام	قدم_زدن_آهسته	راه_رفتن_آهسته	قدم_زدن_آرام	راه_رفتن_با_آرامش	4
		s	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_بی_حوصله	راه_رفتن_اهسته	قدم_زدن_آهسته	4
		rs	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_بافکر	راه_رفتن_ناراحت	راه_رفتن_بی_حوصله	2
limping_slowly	28	-s	لنگان_لنگان_رفتن_بد	لنگان_لنگان_رفتن_آرام	لنگان_لنگان_رفتن_شل	لنگان_لنگان_رفتن_خراب	راه_رفتن_پیر	2
		s	لنگیدن_آهسته	لنگیدن_خسته	لنگیدن_خیلی_آهسته	لنگیدن_ارام	لنگیدن_آرام	4
		rs	لنگیدن_خسته	راه_رفتن_خسته	لنگیدن_آهسته	راه_رفتن_لنگان_لنگان	لنگان_لنگان_رفتن_لنگان	3
limping_painfully	21	-s	لنگان_لنگان_رفتن_بد	لنگان_لنگان_رفتن_شل	لنگان_لنگان_رفتن_خراب	لنگیدن_پادرد	لنگان_لنگان_رفتن_لنگان	3
		s	لنگیدن_خسته	قدم_زدن_بالنگیدن	راه_رفتن_لنگان_لنگان	قدم_برداشتن_اهسته	راه_رفتن_لنگ_لنگان	3
		rs	قدم_زدن_بالنگیدن	راه_رفتن_لنگ_لنگان	راه_رفتن_لنگان_لنگان	راه_رفتن_خسته	راه_رفتن_آرام	3

It is obvious that more appropriate Farsi translations have been extracted for English verb-adverbs. In addition, best translations are ranked more appropriately using pattern-grounding; for instance, the best translation for 'limping_paingfully' is 'لنگیدن_آهسته' which is ranked the 1st using pattern-grounding. On the other hand, it is ranked the third using randomized-grounding approach. Besides that, the other translations found for 'limping_slowly' by random-grouding are more general than the translations found by pattern-grounding. In other words, the translations found by pattern-grounding are more appropriate as the meaning of both verbs and adverbs overlap precisely between the source and target language.

3.4.4. Translation of Farsi annotations to English

Table 21: Translation of the Farsi to English motion verbs

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
راه_رفتن	1637	-s	walking	walk	Walking	Limping	limping	3
		s	walking	walk	Walking	Limping	moving	4
		rs	walking	Limping	limping	Walking	scuffing	2
دویدن	339	-s	running	jogging	Running	Jogging	run	5
		s	running	jogging	Running	Jogging	sprinting	5
		rs	running	jogging	Running	walking	Jogging	4
لنگیدن	312	-s	limping	Limping	scuffing	stagger	hobbling	4
		s	limping	Limping	scuffing	limbing	limb	4
		rs	Limping	limping	scuffing	walking	moving	3

Using pattern-grounding, the translation of Farsi verbs into English verbs is enhanced. In addition, as earlier stated, typos are ignored; thus, 'limb' and 'limbing' are interpreted as 'limp' and 'limping'. In total, 2811 Farsi annotations have been recorded; out of these 2811 annotations, 1637 annotations include the verb 'راه_رفتن'='to walk'. This is because most motions can be described using 'راه_رفتن' combined with an apt adverb. As a result, the Farsi verbs which have been used to annotate the motion videos are much less diverse than the English verbs; in other words, Farsi verbs are more general, while English annotated verbs are more specific. Yet, the translation of these general Farsi verbs to more specific English verbs has good performance.

The most frequent Farsi verb-adverb is 'راه_رفتن_آرام' which is translated into 'to walk slowly'. Although I did not count 'walking_carefully' as one of its good translations, in a deeper point of view, the meaning of 'walking_carefully' is closer to the meaning of 'walking_slowly' than the 'walking_sadly'. It is worth noticing that 'walking_sadly' is rated the third best translation of 'راه_رفتن_آرام' by indexical-grounding approach, while it is ranked down to the 5th best translation. From this point of view, pattern-grounding excels the indexical-grounding.

The second most frequent Farsi verb-adverb is 'راه_رفتن_لنگان_لنگان' which can be best translated to 'walking_limpingly' or 'limping'. Although most of its extracted translations convey the meaning of 'limping', the scope of their meanings is narrowed down by some adverbs; in other words, they are more specific than the verb-adverb in the source language. Hence, none of them is considered as a good translation.

3.4.5. Translation of Farsi annotations to Finnish

Pattern-grounding has a much better performance than the other two approaches in translating 'راه_رفتن'='to walk'. For the other two Farsi verbs, all three approaches has performed fairly well; however, pattern-grounding prevails in translating the less frequent Farsi verbs, while randomized-grounding profoundly deteriorates the translation. Although neither 'kävelee' nor 'raahustaa' has been considered as good translation of 'لنگیدن', the meaning of 'raahustaa'='to scuff' is closer to 'لنگیدن'='to limp'.

Table 22: Translation of the Farsi to English motion verb-modifiers								
Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
راه_رفتن_آرام	389	-s	walking_slowly	walking_very_slowly	walking_sadly	walking_carefully	walking_thoughtfully	2
		s	walking_slowly	walking_carefully	walking_very_slowly	Limping_Normally	walking_sadly	2
		rs	walking_slowly	limping_painfully	Limping_Slowly	Limping_Very_fast	Walking_Slowly	2
راه_رفتن_لنگان_لنگان	346	-s	Limping_Slowly	limping_slowly	Limping_Painfully	walk_injured	limping_painfully	0
		s	limping_painfully	Limping_Slowly	limping_slowly	Limping_Painfully	scuffing_very_slowly	0
		rs	limping_painfully	Limping_Slowly	Limping_Very_fast	walking_slowly	limping_slowly	0
راه_رفتن_کمی_تند	135	-s	Limping_Very_fast	limping_hurriedly	walking_actively	walking_unsteady	walking_slightly_weirdly	1
		s	Limping_Very_fast	Striding_Fast	walking_briskly	walking_heavily	Walking_Quickly	3
		rs	Limping_Very_fast	Limping_Normally	walking_briskly	limping_painfully	limping_hurriedly	1

Table 23: Translation of the Farsi to Finnish motion verbs								
Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
راه_رفتن	1637	-s	kävelee	kävellä	nilkuttaa	ontuu	ontua	2
		s	kävelee	kävellä	löntystelee	Kävelee	nilkuttaa	4
		rs	kävelee	kävellä	nilkuttaa	ontuu	ontua	2
دویدن	339	-s	juosta	juoksee	hölkkää	juokseminen	Juoksee	5
		s	juosta	juoksee	hölkkää	starttaa_juoksuun	juokseminen	5
		rs	juosta	juoksee	hölkkää	hölkyttää	nilkuttaa	4
لنگیدن	312	-s	ontuu	nilkuttaa	ontua	linkuttaa	raahustaa	4
		s	ontuu	nilkuttaa	ontua	linkuttaa	raahustaa	4
		rs	nilkuttaa	ontuu	ontua	kävelee	linkuttaa	4

The most frequent Farsi verb-adverb is 'راه_رفتن_آرام' which can be best translated as 'to walk slowly'. Only the 1st and 2nd extracted translations of the indexical-grounding are accurate, while the the 1st, 4th and 5th extracted translations of the pattern-grounding are good translations. Furthermore, the 3rd extracted Finnish verb-adverb, 'kävelee_kävelee_nilkuttaen_hitaasti', is not considered as a good transaltion because it has ‘nilkuttaen’=’limpingly’ as its additional qualifier adverb which narrows down the meaning of ‘kävelee’=’to walk’.

The second most frequent annotated Farsi verb-adverb is ‘راه_رفتن_لنگان_لنگان’=’to walk limpingly’. It is a general word covering the meaning of ‘limping’ action. However, the only extracted translation which is as general is ‘kävelee_ontuen’. All the other extracted translations have more specific meaning. Since it is not accurate to translate from a general word to a specific interpretation of that word, I have counted only one good translation for ‘راه_رفتن_لنگان_لنگان’. Besides that, the strength of pattern-grounding is demonstrated in the accurate translation of the third most frequent Farsi verb-adverb. All its 5 extracted translations are apt and precise.

Table 24: Translation of the Farsi to Finnish motion verb-modifiers

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
راه_رفتن_آرام	389	-s	kävelee_hitaasti	kävellä_hitaasti	kävelee_mietteliäästi	kävelee_varovasti	kävelee_surullisena	2
		s	kävelee_hitaasti	kävelee_epävarmasti	kävelee_kävelee_nilkuttaen_hitaasti	kävelee_rauhallisesti	kävelee_hyvin_hitaasti	3
		rs	kävelee_hitaasti	ontuu_hitaasti	kävelee_rammiten	kävelee_ontuen	kävelee_rennosti	1
راه_رفتن_لنگان_لنگان	346	-s	kävelee_ontuen	ontuu_kivuliaasti	ontuu_hitaasti	ontuu_vaivalloisesti	nilkuttaa_hitaasti	1
		s	ontuu_pahasti	kävelee_ontuen	ontuu_hitaasti	ontuu_kivuliaasti	kävelee_rammiten	1
		rs	kävelee_ontuen	ontuu_hitaasti	ontuu_hyvin_kivuliaasti	ontuu_pahasti	kävelee_rammiten	1
راه_رفتن_کمی_تند	135	-s	kävellä_reipas	kävelee_tomerasti	kävelee_ontuen	kävelee_hiukan_ontuen	kävelee_reippaasti	3
		s	kävelee_reippaasti	kävelee_tomerasti	kävelee_nopeasti	kävelee_tarmokkaasti	kävellä_reippaasti	5
		rs	kävelee_ontuen	ontuu_pahasti	kävellä_reipas	kävelee_rammiten	kävelee_reippaasti	2

3.4.6. Translation of Finnish annotations to Farsi

In Table 25, we can see that Finnish verbs have been translated properly to Farsi verbs because they are more specific expression of motions than the Farsi annotated verbs;

Table 26 demonstrates that the translation of Finnish annotated verb-adverbs to Farsi verb-adverbs has been boosted by applying pattern-grounding. The translations extracted by pattern-grounding approach are more precise because the meaning of both verb and adverb overlap significantly. For example, using randomized-grounding and indexical-grounding, only two translations out of the extracted translations for ‘ontuu_hitaasti’ covers both the meaning of ‘ontuu’ and ‘hitaasti’, while using pattern-grounding, four translations cover precisely the meaning of ‘ontuu’ and ‘hitaasti’.

Table 25: Translation of the Finnish to Farsi motion verbs								
Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	#good translation
kävelee	618	-s	راه_رفتن	راهرفتن	قدم_زدن	گام_برداشتن	تند_راه_رفتن	4
		s	راه_رفتن	راهرفتن	قدم_زدن	گام_برداشتن	لنگان_لنگان_راه رفتن	4
		rs	راه_رفتن	قدم_زدن	گام_برداشتن	راهرفتن	لنگان_لنگان_راه رفتن	4
ontuu	229	-s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_ر فتن	راه_رفتن	لنگان_لنگان_راه _رفتن	5
		s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_ر فتن	کشال_کشال_را ه_رفتن	قدم_برداشتن	5
		rs	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	قدم_زدن	گام_برداشتن	5
nilkuttaa	164	-s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_ر فتن	راه_رفتن	می_لنگد	5
		s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_ر فتن	قدم_برداشتن	لنگان_لنگان_راه _رفتن	5
		rs	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	قدم_زدن	گام_برداشتن	5

3.4.7. All translations results

Table **27** demonstrates that verbs can be translated more efficiently than when verbs are combined with adverbs. This is because verbs have been used to annotate a set only three different motions {walking, running, or limping}, while the selection of adverbs depend on how an annotator observe a motion; in other words, the same motion can be described using various adverbs, and the scope of the variety of used adverbs is as large as the view of a person. In addition, the adverbs explain the style of a motion, and an annotator view the style of a motion from a different angle. This means the selection of adverbs is culture dependent. As a result, matching these diverse adverbs between two different languages would be more difficult than matching verbs.

Our observation shows that pattern-grounding is helpful in translating verbs. When pattern-grounding is applied, the good translations will get higher rank than the irrelevant extracted translations. Furthermore, since randomized-grounding has worsened the translation of verbs significantly, the application of pattern-grounding in translating the motion verbs seems to be a meaningful and a reasonable process.

The translation of verb-adverbs has also been enhanced when pattern-grounding is exploited. However, the translation of English and Finnish verb-adverbs to Farsi verb-adverbs was not deteriorated by randomized-grounding. It is, nevertheless, worth mentioning that in all our experiments, pattern-grounding surpasses randomized-grounding in translating verbs and verb-adverbs.

Table 26: Translation of the Finnish to Farsi motion verb-modifiers

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
kävelee_hitaasti	64	-s	راه_رفتن_متفکرانه	قدم_زدن_آهسته	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_بی_حوصله	3
		s	راه_رفتن_متفکرانه	راه_رفتن_بی_حوصله	قدم_زدن_آهسته	راه_رفتن_آهسته	راه_رفتن_بی_هدف	2
		rs	راه_رفتن_متفکرانه	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_بی_حوصله	2
kävelee_reippaasti	52	-s	راه_رفتن_مصمم	راه_رفتن_تند	راه_رفتن_سریع	راه_رفتن_مغرور	راهرفتن_به_طوری_نرمال	2
		s	راه_رفتن_مصمم	راه_رفتن_خوشحال	راه_رفتن_با_عجله	راه_رفتن_با_انرژی	راه_رفتن_تند	3
		rs	راه_رفتن_سریع	راه_رفتن_تند	راه_رفتن_با_عجله	راه_رفتن_مصمم	راه_رفتن_خوشحال	3
ontuu_hitaasti	47	-s	لنگیدن_خسته	لنگان_لنگان_رفتن_بد	لنگیدن_آهسته	راه_رفتن_لنگان_لنگان	لنگان_لنگان_رفتن_آرام	3
		s	لنگیدن_آهسته	لنگیدن_خسته	لنگیدن_آرام	لنگیدن_خیلی_آهسته	لنگیدن_به_آرامی	4
		rs	لنگیدن_خسته	لنگیدن_آهسته	راه_رفتن_لنگان_لنگان	راه_رفتن_خسته	راه_رفتن_آرام	2

Table 27: All translation results

	Number of good extracted translations (out of 50)		
	indexical grounding	pattern grounding	randomized grounding
English verbs to Farsi verbs	35	37	29
Farsi verbs to English verbs	39	40	26
Farsi verbs to Finnish verbs	39	40	25
Finnish verbs to Farsi verbs	44	43	39
Finnish verbs to English verbs	38	42	32
English verbs to Finnish verbs	43	44	24
English verb-adverbs to Finnish verb-adverbs	17	23	23
Finnish verb-adverbs to English verb-adverbs	23	24	17
Farsi verb-adverbs to English verb-adverbs	19	21	11
English verb-adverbs to Farsi verb-adverbs	22	27	22
Farsi verb-adverbs to Finnish verb-adverbs	24	26	15
Finnish verb-adverbs to Farsi verb-adverbs	27	36	27

3.5. Hierarchical clustering result

Farsi verbs, Farsi verb-adverbs, English verbs, English verb-adverbs, Finnish verbs, and Finnish verb-adverbs have been clustered using agglomerative hierarchical clustering in both indexical-grounding and pattern-grounding cases. The cophenetic correlation value has been computed for the resulting dendrograms. Ward, median, and centroid methods have not been considered because cosine distance is employed as a metric to compute the similarity of objects. In the following section, the hierarchical clustering result of the average method is displayed.

3.5.1. Hierarchical result of English annotations

In this section, you can see the dendrograms which have been resulted from employing the agglomerative hierarchical clustering on English verbs which have been used to annotate the motion videos. Besides that, the resulting dendrogram has been cut horizontally to produce some clusters.

Figure 13 shows the dendrogram achieved by applying agglomerative hierarchical clustering on English annotated verbs. Only the 30 most frequent English annotated verbs have been included in the dendrogram. These verbs are not normalized by the pattern-grounding process; in other words, the dendrogram displays the result of indexical-grounding approach. Furthermore, the average method is exploited to compute the similarity of objects and clusters.

Figure 14 displays the dendrogram of pattern-grounded English verbs. As stated in section 2.4.2.1, a dendrogram can be cut horizontally into any arbitrary number of clusters. This makes dendrogrogram a very useful information vizualization tool; a researcher can observe a dendrogram and decide subsequently about the number of clusters. I have decided to cut the dendrogram in Figure 14 into 4 clusters; in order to be fair, the indexical-grounding dendrogram in Figure 14 is also cut into 4 clusters.

	Table 28: hierarchical clustering result of English verbs
Indexical-grounding	cluster1={amble, Striding} cluster2={walking, limping, Limping, Walking, scuffing, marching, walk, stomping, shuffling, moving, strolling, limb, wandering, Strolling, shuffle, hobbling, stagger, meandering, striding, plodding, leaping, dancing} cluster3={limbing} cluster4={running, jogging, Running, Jogging, sprinting}
Pattern-grounding	cluster1={marching, stomping, dancing} cluster2={running, jogging, Running, Jogging, sprinting} cluster3={limping, Limping, scuffing, shuffling, moving, limb, limbing, shuffle, hobbling, stagger, plodding, leaping} cluster4={walking, Walking, walk, strolling, wandering, Strolling, amble, Striding, meandering, striding}

Using pattern-grounding, the English verbs are clustered appropriately, while using indexical-grounding approach, they spread erratically into inappropariate clusters. For example, the verbs which explain walking and limping are mixed in cluster2, while they are separated into two suitable clusters using pattern-grounding. In addition, cluster3 contains only 'limb' as its single member. On the other hand, when English verbs are pattern-grounded, they go to their correct and precise cluster. Cluster1 represent the 'marching' motion; cluster2 denotes 'running; cluster3 exemplify 'limping'; and cluster4 typify 'walking'. The only verb which has been misclassified is 'leaping'.

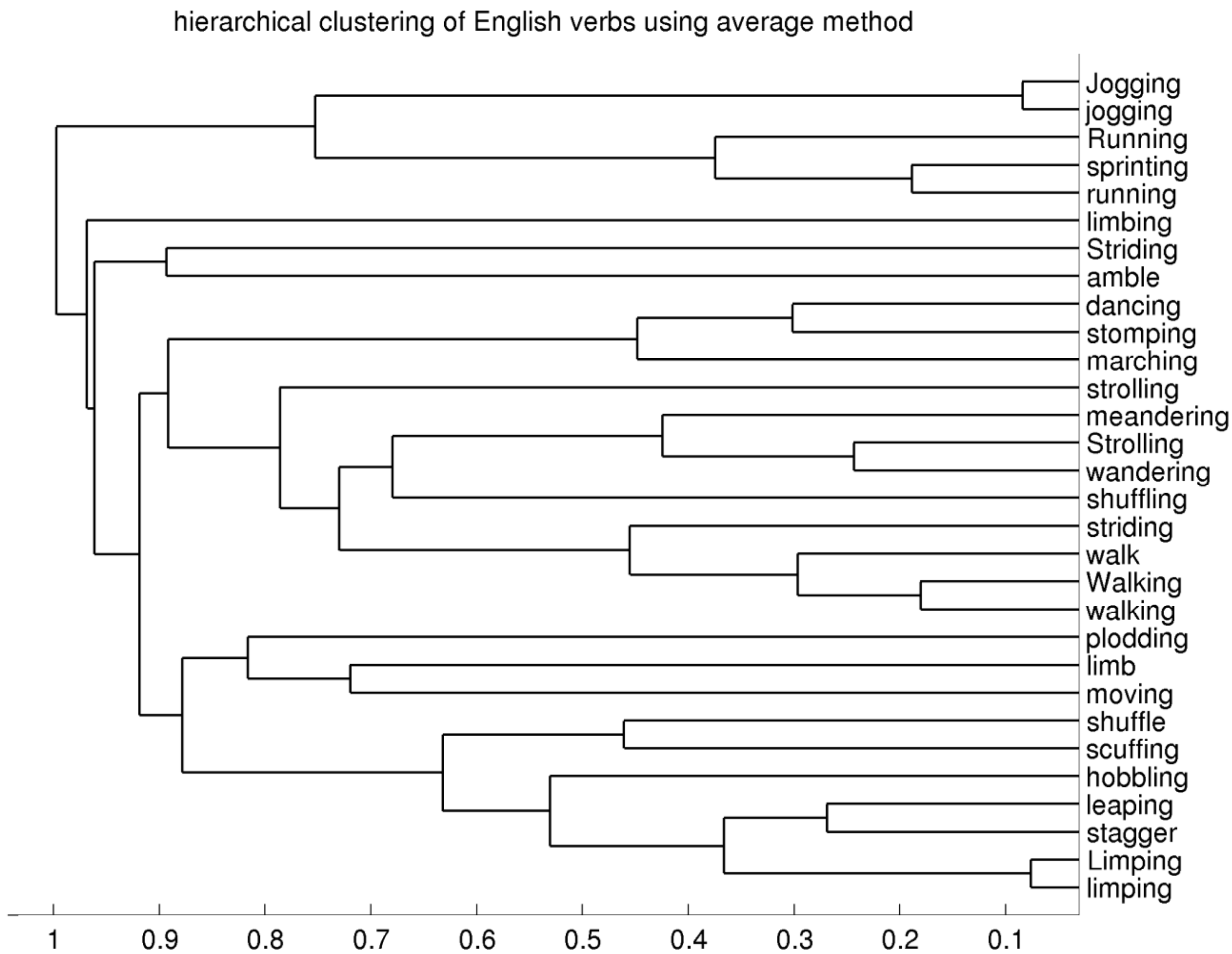


Figure 13: Hierarchical clustering of the English motion verbs using **Indexical-grounding** approach. The above figure demonstrates the dendrogram obtained by applying agglomerative hierarchical clustering on 30 most frequent English verbs which have been used to annotate the motion videos. The average method has been utilized to compute the distance between two clusters.

In addition to correctly clustering the English verbs, the pattern-grounded dendrogram looks more consistent than the indexical-grounded dendrogram. This property can be observed by comparing the height of the links in the dendrograms. This observation can be made more precise by computing the cophenetic correlation coefficient. The cophenetic coefficient of the indexical-grounding dendrogram is 0.87, and it is 0.89 for the pattern-grounding case.

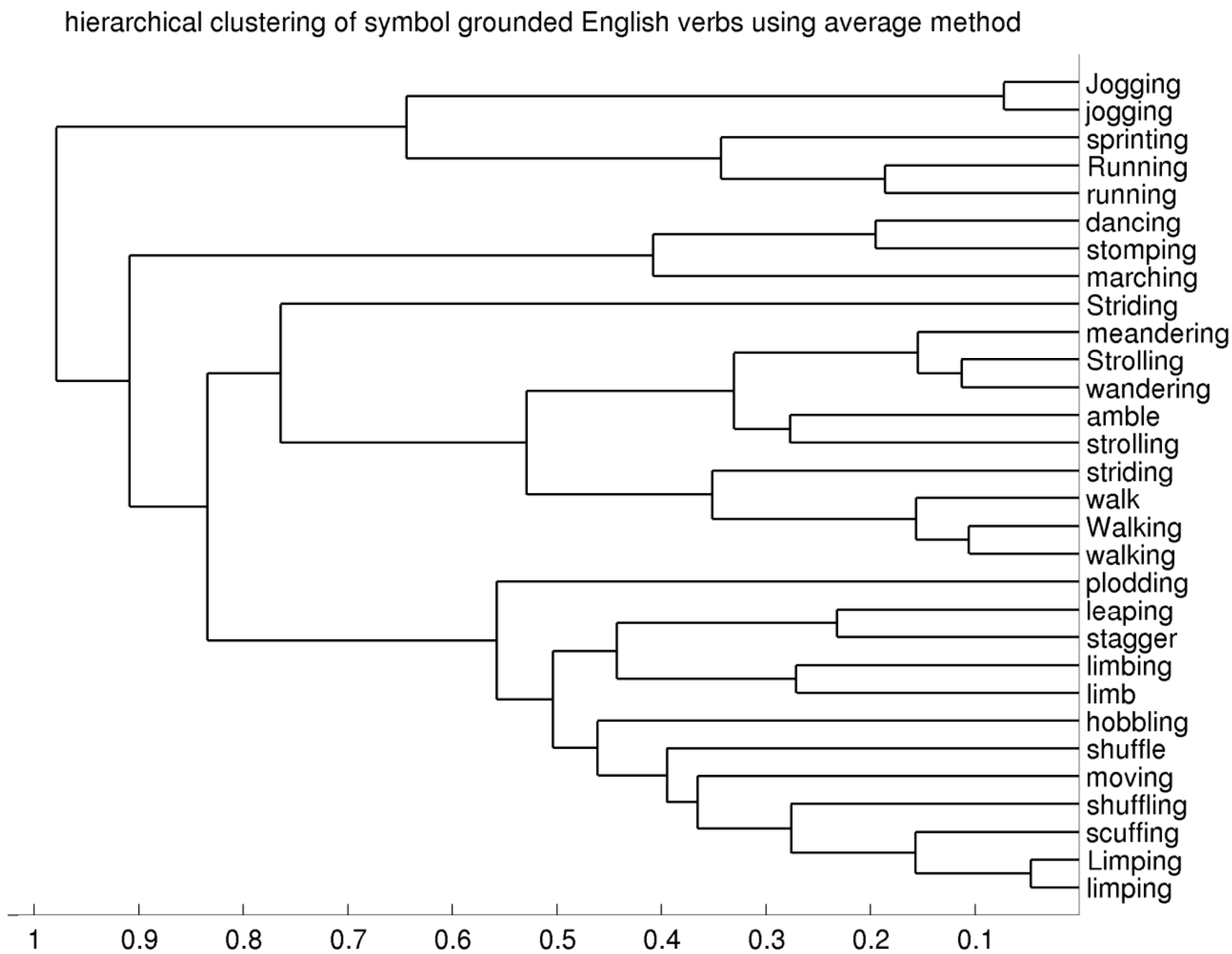


Figure 14: Hierarchical clustering of the English motion verbs using **pattern grounding** approach.The above figure displays the dendrogram of English verbs. The average method has been used to determine the distance between objects and clusters pattern-grounded. In addition, only the 30 most frequent English verbs have been included in the dendrogram.

3.5.2. Hierarchical clustering result of Finnish verbs

Finnish verbs have been represented by 124-dimensional vectors. There is one component for every video. Originally, it is the frequency of a verb in a video which is stored in the corresponding component; next, using the pattern-grounding process, the components of these vectors are normalized. These vectors have been hierarchically clustered using average, complete, single, and weighted methods. Figure 15 and Figure 16 demonstrates the dendrogram obtained by the average method.

Both of the dendrograms in Figure 15 and Figure 16 have been cut vertically into 4 clusters. The resulting clusters are included in the following table.

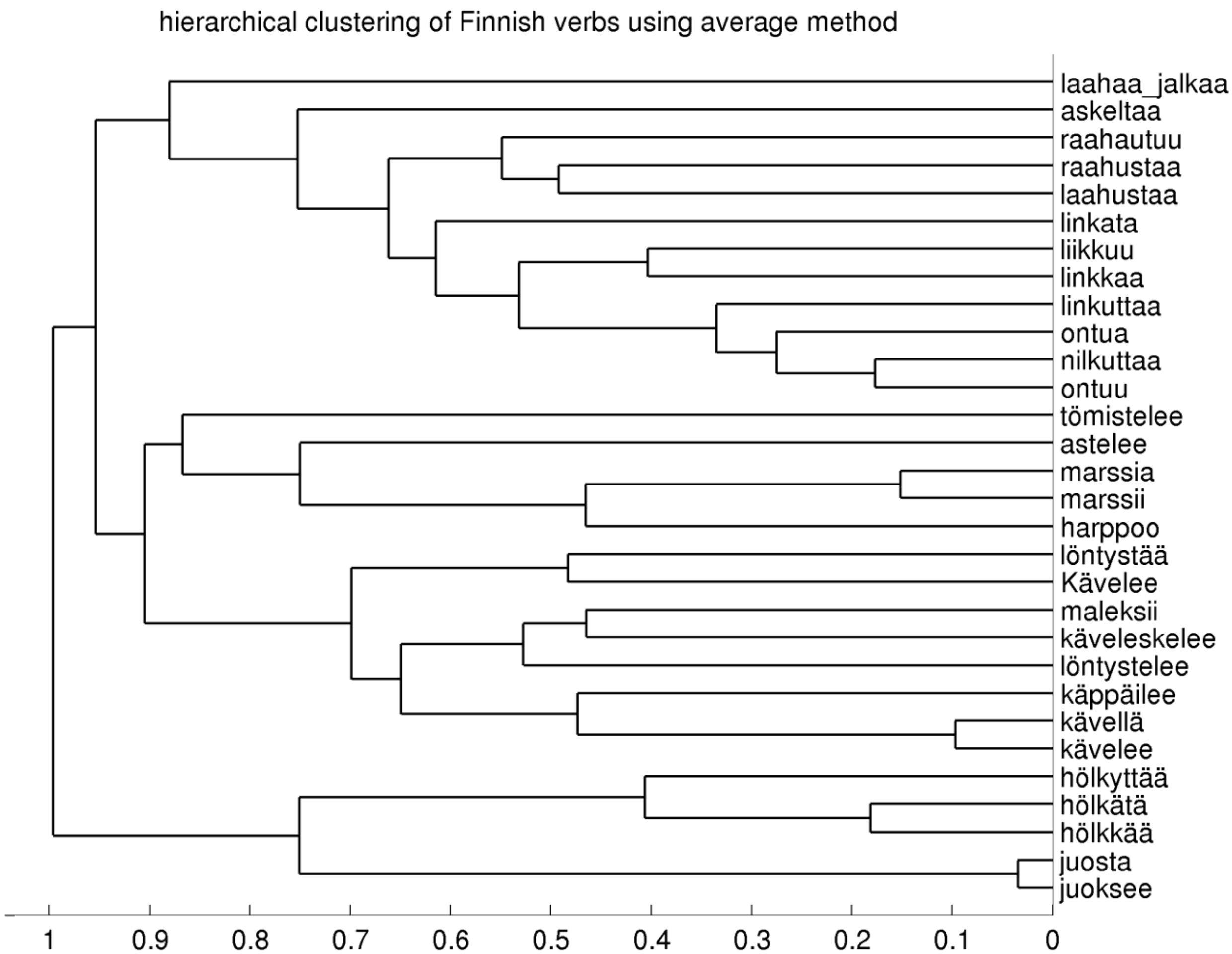


Figure 15: Agglomerative hierarchical clustering of Finnish verbs using the **indexical-grounding** approach. The verbs are indexically grounded. The average method has been selected as the similarity measure.

	Table 29: hierarchical clustering result of Finnish verbs
Indexical-grounding	cluster1={kävelee, kävellä, löntystelee, Kävelee, käveleskelee, maleksii, käppäilee, löntystää} cluster2={harppoo, marssii, tömistelee, astelee, marssia } cluster3={ ontuu, nilkuttaa, ontua, linkuttaa, laahustaa, raahustaa, raahautuu, linkkaa, askeltaa, laahaa_jalkaa, linkata, liikkuu } cluster4={ juoksee, hölkkää, juosta, hölkyttää, hölkätä}
Pattern-grounding	cluster1={kävelee, kävellä, löntystelee, Kävelee, käveleskelee, astelee, maleksii, käppäilee, löntystää} cluster2={ ontuu, nilkuttaa, ontua, linkuttaa, laahustaa, raahustaa, raahautuu, linkkaa, askeltaa, laahaa_jalkaa, linkata, liikkuu } cluster3={ harppoo, marssii, tömistelee, marssia} cluster4={ juoksee, hölkkää, juosta, hölkyttää, hölkätä}

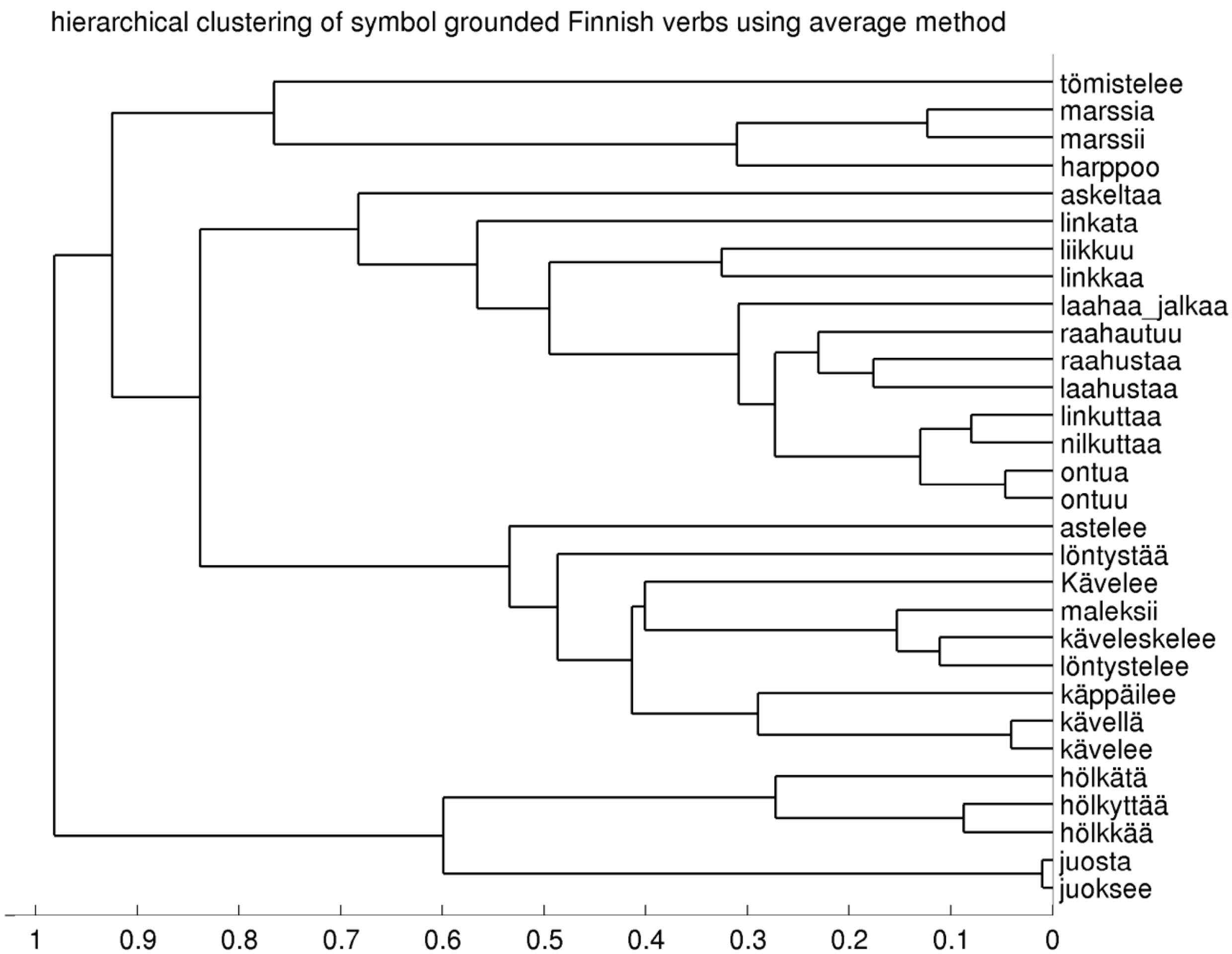


Figure 16: The agglomerative hierarchical clustering of pattern-grounded Finnish verbs using the **pattern-grounding** approach. The average method has been chosen as the similarity measure.

Both indexical-grounding and pattern-grounding approaches divided the Finnish verbs into appropriate clusters. In fact, they only difference between the clustering results of these two approaches is that the verb 'astelee' is assigned to cluster2 being recognized as the synonym of 'harpoo', while using pattern-grounding, it is assigned to cluster1 being identified as a synonym of 'kävelee'. Besides that, the cophenetic correlation coefficient has been improved from 0.91 in indexical-grounding case to 0.93 in pattern-grounding case. In other words, the pattern-grounded dendrogram is more consistent than the indexical-grounded dendrogram.

3.5.3. Hierarchical clustering of Farsi annotations

Farsi verbs, which have been used to annotate motion videos, have also been clustered using aggolomerative hierarchical clustering. The dendrogram resulting from this hierarchical clustering is cut vertically into 4 communities.

Table 30 shows the resulting clusters when the dendrogram in Figure 17 and Figure 18 are cut vertically into 4 slices.

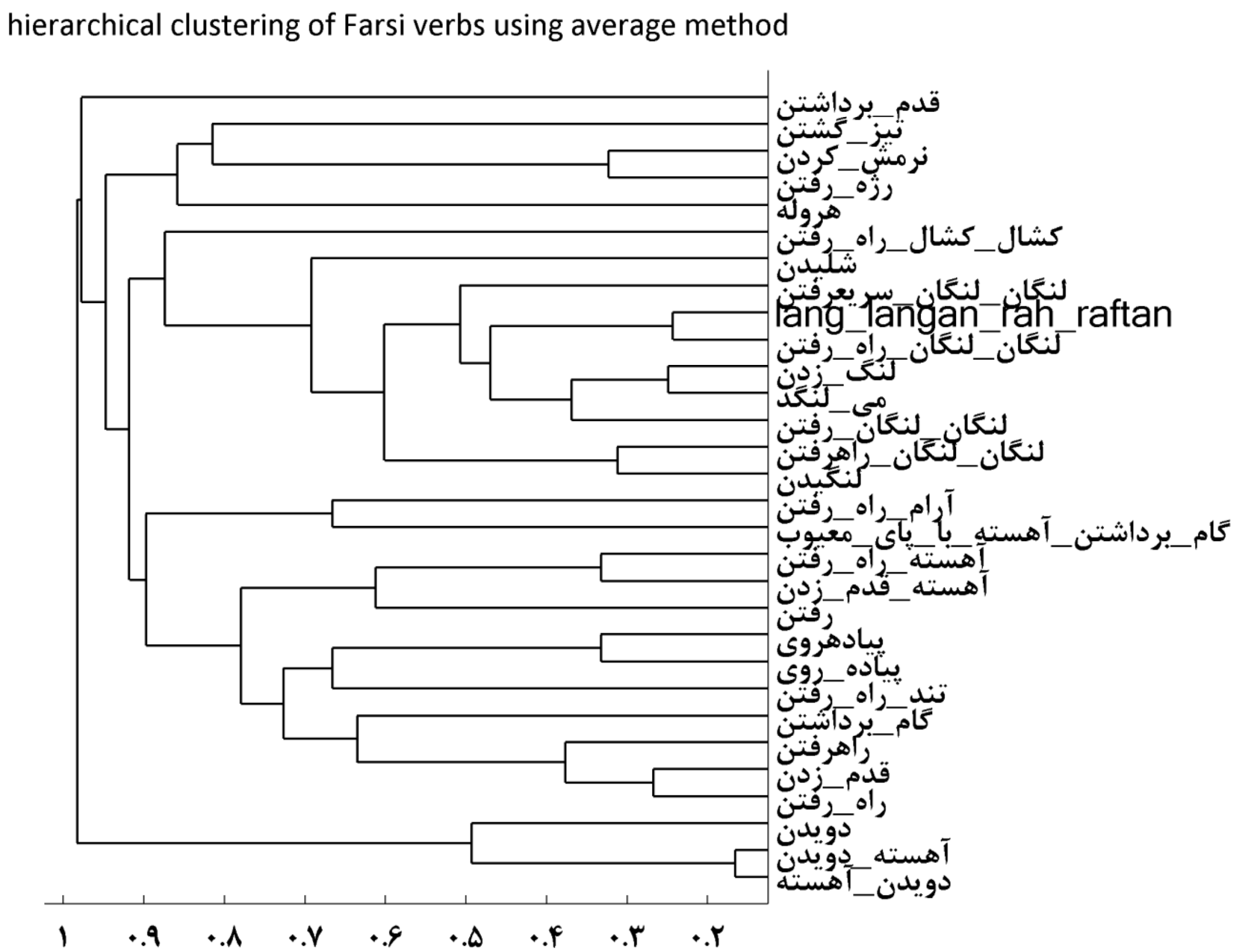


Figure 17: Agglomerative hierarchical clustering of indexical-grounded Farsi verbs using the **indexical-grounding** approach. Average method is selected as the similarity measure.

	Table 30: hierarchical clustering result of Farsi verbs
Indexical-grounding	cluster1={ هروله, رژه_رفتن, نرمش_کردن, تیز_گشتن} cluster2={ قدم_زدن,لنگیدن,راه_رفتن,می_لنگد,رفتن,لنگان_لنگان_راه_رفتن,راهرفتن,لنگان_لنگان_رفتن,گام_برداشتن,لنگان_لنگان_راهرفتن,لنگان_لنگان_سریعرفتن, lang_langan_rah_raftan,کشال_کشال_راه_رفتن,لنگ_زدن,آهسته_قدم_زدن,پیادهروی,پیاده_روی,تند_راه_رفتن,شلیدن,گام_برداشتن_آهسته_با_پای_معیوب,آرام_راه_رفتن, آهسته_راه_رفتن} cluster3={قدم_برداشتن} cluster4={آهسته_دویدن,دویدن_آهسته,دویدن}
Pattern-grounding	cluster1={پیادهروی,پیاده_روی,تند_راه_رفتن,تیز_گشتن,نرمش_کردن,رژه_رفتن,گام_برداشتن} cluster2={آرام_راه_رفتن,آهسته_راه_رفتن,آهسته_قدم_زدن,رفتن,راهرفتن,قدم_زدن,راه_رفتن} cluster3={می_لنگد,لنگان_لنگان_راه_رفتن,قدم_برداشتن,لنگان_لنگان_رفتن,لنگان_لنگان_راهرفتن,لنگیدن,لنگان_لنگان_سریعرفتن,lang_langan_rah_raftan,کشال_کشال_راه_رفتن,لنگ_زدن,گام_برداشتن_آهسته_با_پای_معیوب,شلیدن} cluster4={آهسته_دویدن,دویدن_آهسته,هروله,دویدن}

Let’s first analyze the clusters resulted from cutting the indexical-grounded dendrogram.Cluster1 of this method has 4 members with every verb being a distinct motion; in other words, the verbs which belong to this cluster are not synonym of each other. Cluster2 consists of the synonyms of both ‘لنگیدن’=’limping’ and ‘راه_رفتن’=’walking’. Cluster3 has only one member. Cluster4 has indeed relevant members; they are all synonyms of ‘دویدن’=’running’. The main property of a good clustering is that the members of a cluster consist with the cluster from some point of view; since we are interested in dividing the verbs, one expect that the members of a cluster be synonyms of a specific verb. In addition, in a good clustering, the clusters have comparable sizes; one would not see neither too large nor too small communities. Since the verbs in cluster1 and cluster2 are not synonym of a single verb, and because cluster3 has only one member, the indexical-grounded dendrogram is not considered to be a desirable hierarchical dendrogram. On the pther hand, the pattern-grounded hierarchical dendrogram has led to perfect clusters both from the semantic and from the size point of view. Cluster1 consists of synonyms of ‘رژه_رفتن’=’marching’; cluster2 contains synonyms of ‘راه_رفتن’=’walking’; cluster3 includes synonyms of ‘لنگیدن’=’limping’; and, cluster4 cover synonyms of ‘دویدن’=’running’; also, notice that cluster2 and cluster1 which are related comes one after another.

hierarchical clustering of symbol-grounded Farsi verbs using average method

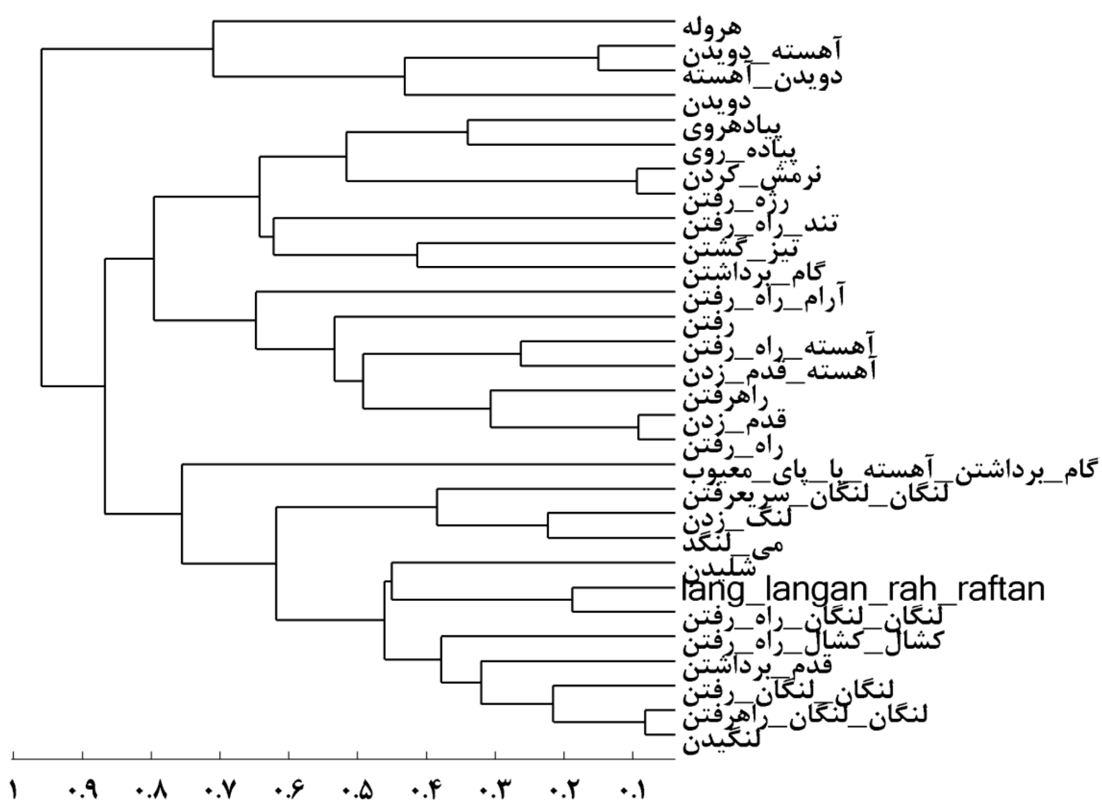


Figure 18: Agglomerative hierarchical clustering of pattern-grounded Farsi verbs using the **pattern-grounding** approach. Average method is chosen as the similarity measure.

3.5.4. Cophenetic correlation coefficients

In Table **31**, I have just reported the average method results; if you are interested in other methods, you can see the tables in the appendix. Table **31** shows that using pattern grounding a dendrogram is created that represents the original distances slightly better than when symbols are just indexically grounded. In other words, when verbs or verb-adverbs are grounded using visual data, the resulting dendrograms are more consistent. The only case which the cophenetic correlation value decreases is in clustering Farsi pattern-grounded verbs. However, as we have seen in the section 3.5.3, pattern-grounding leads to accurate clustering of Farsi verbs. There might be a random normalization which even excels the cophenetic value of the pattern-grounding result; however, according to the observation we had, there is no guarantee that such randomized-grounding would produce meaningful cluster.

Table 31: cophenetic values of the agglomerative hierarchical clustering using the average method

	Cophenetic value	
Dataset	indexical grounding	pattern grounding
Farsi verbs	0.8150	0.7918
Farsi verb-adverbs	0.9070	0.9443
English verbs	0.8722	0.8921
English verb-adverbs	0.8832	0.8935
Finnish Verbs	0.9123	0.9313
Finnish verb-adverbs	0.8645	0.9162

4. Discussion and Conclusion

The application of ICA on word-document matrix extracted from Tabnak and Alef corpora led to the detection of syntagmatic word clusters such as {Quran, God, religion}, {oil, nuclear (power), Iran, inter, national}, {Gaza, Zionist (Regime), USA}, and {currency, Dollar}. On the other hand, utilization of ICA on the word-word matrices, extracted from these two corpora, resulted in the discovery of paradigmatic word clusters such as synonyms in addition to the syntagmatic word clusters. The resulting clusters of both of these two methods can be exploited in the automatic construction of a thesaurus.

Both individuals and search engines can benefit from this kind of automatic thesaurus. Such thesaurus can be viewed as the summary of a large corpus; for example, the extracted word clusters of these two corpora were mainly about politics, sport, religion and war. In addition, some of the word clusters can shed light on the strategic planning of a news agency; for instance, the emergence of these terms in one cluster {Sepah, culture, commander, language, (Imam) Hussein} can indicate the exploitation of religion for both political and cultural ends.

Such automatic thesaurus can also assist an individual to overcome his vocabulary problem; a person may know what he is looking for, but he is unable to articulate the problem in terms recognized by the search engine. For instance, a user who is searching for “Dollar” in the context of Iranian economy is very likely to be also interested in reading about “currency” related topics. Furthermore, it can serve as a brainstorming tool. For example, a researcher who is investigating about ‘oil’ might be enlightened if he also studies about ‘nuclear power’ as these two key terms reveal a decisive relationship between Iran and the international community.

Semantic-based search engines could be reinforced by using such automatic thesaurus. The relevance of a document can be determined by the frequency of a key term and the utilization of semantically related words. For example, if the search term is “Quran”, the documents which contains the terms “God” and “religion” would be more inclusive than documents containing only “Quran”; thus, they can obtain a higher ranking placement in the list of search results. Besides that, a search engine can employ such automatic thesaurus for enhancing its keyword suggestion tool. For instance, if a user is seeking for “Gaza”, he can be suggested to also search for “Zionist Regime” and “USA”. Of course, all of the aforementioned applications of an automatic thesaurus extracted by ICA might be over-ambitious ideas as some of the extracted words are not so strongly and semantically related.

For the second part of the project, motion data was analyzed. In order to see whether symbol grounding can be effective in the detection of synonyms, we have implemented an experiment. First of all, every motion video was transformed into a 602-dimensional vector using the motion data. The motion data includes features such as the means and standard deviations of coordinates, velocities, and accelerations of different body parts. After that, the dimensionality of the motion data was reduced using PCA. Besides that, every video was portrayed as a pie in a 2-dimensional space using the first two principal components as the coordinates of the pies.

According to our observation, there was a large variation in how people used “راه رفتن” =“walking”. Furthermore, people were more unanimous about verbs than modifiers. The most interesting part of this experiment was that the three main locomotion verbs namely {‘walking’, ‘راه-رفتن’ , ‘kävelee’}, {‘limping’, ‘لنگیدن’, ‘ontuu’} and {‘running’, ‘دویدن’, ‘juoksee’} occupied roughly the same space. This propelled us to transform verbs and modifiers into vectors by using video indexes as their context. We called this type of transformation as the indexical-grounding.

Another observation was that synonyms emerge close to each other. In other words, the verbs and verb-adverbs in similar videos tend to be semantically related. The video similarity could be calculated using the motion data. Using the motion data and a threshold, we defined a neighborhood for every video. A video influenced all of its neighbors by normalizing them according to their distances. In order to understand this normalization, we should note that every video is associated with two vectors. One vector is calculated based on the raw mocap data obtained by Arena; one can extract the video similarities from this vector. The second vector includes the frequencies of verbs and verb-adverbs; there was one component for every verb and verb-adverb. A portion of the vector components of every video have been added to the vector components of the neighboring videos. Since the vectors corresponding the verbs and verb-adverbs have been normalized by the motion data, textual and visual modalities have been fused. We referred to this type of fusion as the pattern-grounding as it was based on the content of the videos.

In addition to normalizing using the original motion data, verb and verb-modifier vectors were also normalized using a random data; this normalization was referred to as randomized-grounding. We compared these three cases of grounding in terms of synonym detection and translation performance.

Although both pattern-grounding and indexical-grounding methods performed well in terms of the number of good detected synonyms, in the thesaurus extracted by the pattern-grounding method, the appropriate synonyms got a higher ranking placement. On the other hand, the thesaurus extraction performance deteriorated significantly using randomized-grounding, which indicates that the selected motion features were apt and felicitous. Furthermore, Pattern-grounding enhanced the translation slightly by placing the more appropriate translations to a higher ranking than the indexical-grounding.

Hierarchical clustering was also applied to extract four clusters from the dendrogram of the 30 most frequent Finnish, Farsi and English annotated motion verbs in both the indexical and the pattern-grounding cases. When pattern-grounding was employed, the clusters were balanced in terms of the number of belonging members; every member of these clusters fits semantically to the other members of the same cluster; and the cophenetic correlation coefficient enhanced slightly. On the other hand, by utilizing the indexical-grounding, some of the extracted clusters had only one member, and the semantically related verbs dispersed arbitrarily to different clusters.

All things considered, we proposed a simple but novel method to fuse textual and visual modalities; we also observe that such fusion has been a useful process in synonym detection and translation. It might be too soon to claim that we have found an anchor or a way for connecting the verbs to their corresponding physical action, but it is safe to say that a at least a glimmer of light from the practical perspective has been shed on the so-called “symbol grounding problem”.

The 602-dimensional variables, which were calculated based on the raw data coming from the Arena software, were the representative of every one of the 124 motion videos. It might be possible to check which features would be most informative to classify or cluster the motion verbs in each language. For instance, the ankle’s position and speed might be adequate to recognize “walking” motion. Based on this inductive knowledge one can experiment whether it would be ever possible to map from selected features representing physical motions to their corresponding verbs and modifiers. One can apply supervised approaches to learn such mapping from numerical representatives of motions to verbs and modifiers.

5. Bibliography

[1] Bell AJ, Sejnowski TJ (November 1995). "An information-maximization approach to blind separation and blind deconvolution". Neural Comput 7 (6): 1129–59.

[2] Bell AJ, Sejnowski TJ (December 1997). "The "Independent Components" of Natural Scenes are Edge Filters". Vision Res. 37 (23): 3327–38.

[3] Hyvärinen, A.; Oja, E. (2000). "Independent component analysis: Algorithms and applications". Neural Networks 13 (4–5): 411–430.

[4] Hyvarinen, A. (1999). "Fast and robust fixed-point algorithms for independent component analysis". IEEE Transactions on Neural Networks 10 (3): 626–634.

[5] A. Cichocki and S.-I. Amari, Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. John Wiley & Sons, Inc., 2002.

[6] A. Hyvarinen, J. Karhunen, and E. Oja, Independent Component Analysis. Wiley-Interscience, May 2001.

[7] A. J. Bell and T. J. Sejnowski, “The "independent components" of natural scenes are edge filters.” Vision Res, vol. 37, no. 23, pp. 3327–3338, December 1997.

[8] O. Déniz, M. Castrillón, and M. Hernández, “Face recognition using independent component analysis and support vector machines,” Pattern Recogn. Lett., vol. 24, no. 13, pp. 2153–2157, 2003.

[9] S. Fiori, “Overview of independent component analysis technique with an application to synthetic aperture radar (sar) imagery processing,” Neural Netw., vol. 16, no. 3-4, pp. 453–467, 2003.

[10] H. Wang, Y. Pi, G. Liu, and H. Chen, “Applications of ica for the enhancement and classification of polarimetric sar images,” Int. J. Remote Sens., vol. 29, no. 6, pp. 1649–1663, 2008.

[11] M. S. Karoui, Y. Deville, S. Hosseini, A. Ouamri, and D. Ducrot, “Improvement of remote sensing multispectral image classification by using independent component analysis,” in 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing. IEEE, August 2009, pp. 1–4.

[12] L. Xiaochun and C. Jing, “An algorithm of image fusion based on ica and change detection,” in Proceedings 7th International Conference on Signal Processing, 2004. Proceedings. ICSP ’04. 2004. IEEE, 2004, pp. 1096–1098.

[13] R. Llinares and J. Igual, “Application of constrained independent component analysis algorithms in electrocardiogram arrhythmias,” Artif. Intell. Med., vol. 47, no. 2, pp. 121–133, 2009.

[14] E. Niedermeyer and F. L. Da Silva, Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams and Wilkins; 4th edition, January 1999.

[15] J. C. Rajapakse, A. Cichocki, and Sanchez, “Independent component analysis and beyond in brain imaging: Eeg, meg, fmri, and pet,” in Neural Information Processing, 2002. ICONIP ’02. Proceedings of the 9th International Conference on, vol. 1, 2002, pp. 404–412 vol.1.

[16] J. Wisbeck, A. Barros, and R. Ojeda, "Application of ica in the separation of breathing artifacts in ecg signals," 1998.

[17] M. Scherg and D. Von Cramon, "Two bilateral sources of the late aep as identified by a spatiotemporal dipole model." Electroencephalogr Clin Neurophysiol, vol. 62, no. 1, pp. 32–44, January 1985.

[18] R. Phlypo, V. Zarzoso, P. Comon, Y. D'Asseler, and I. Lemahieu, "Extraction of atrial activity from the ecg by spectrally constrained ica based on kurtosis sign," in ICA'07: Proceedings of the 7th international conference on Independent component analysis and signal separation. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 641–648.

[19] J. Oster, O. Pietquin, R. Abächerli, M. Kraemer, and J. Felblinger, "Independent component analysis-based artefact reduction: application to the electrocardiogram for improved magnetic resonance imaging triggering," Physiological Measurement, vol. 30, no. 12, pp. 1381–1397, December 2009. [Online]. Available: http://dx.doi.org/10.1088/0967-3334/30/12/007

[20] R. Vigário, J. Särelä, V. Jousmäki, M. Hämäläinen, and E. Oja, "Independent component approach to the analysis of eeg and meg recordings." IEEE transactions on bio-medical engineering, vol. 47, no. 5, pp. 589–593, May 2000.

[21] J. Onton, M. Westerfield, J. Townsend, and S. Makeig, "Imaging human eeg dynamics using independent component analysis," Neuroscience & Biobehavioral Reviews, vol. 30, no. 6, pp. 808–822, 2006.

[22] B. Jervis, S. Belal, K. Camilleri, T. Cassar, C. Bigan, D. E. J. Linden, K. Michalopoulos, M. Zervakis, M. Besleaga, S. Fabri, and J. Muscat, "The independent components of auditory p300 and cnv evoked potentials derived from single-trial recordings," Physiological Measurement, vol. 28, no. 8, pp. 745–771, August 2007.

[23] J. C. Mosher, P. S. Lewis, and R. M. Leahy, "Multiple dipole modeling and localization from spatiotemporal meg data," Biomedical Engineering, IEEE Transactions on, vol. 39, no. 6, pp. 541–557, 1992.

[24] A. C. Tang and B. A. Pearlmutter, "Independent components of magnetoencephalography: localization," pp. 129–162, 2003.

[25] J. Parra, S. N. Kalitzin, and Lopes, "Magnetoencephalography: an investigational tool or a routine clinical technique?" Epilepsy & Behavior, vol. 5, no. 3, pp. 277–285, June 2004.

[26] K. Petersen, L. K. Hansen, T. Kolenda, and E. Rostrup, "On the independent components of functional neuroimages," in Third International Conference on Independent Component Analysis and Blind Source Separation, 2000, pp. 615–620.

[27] C.-J. Lu, T.-S. Lee, and C.-C. Chiu, "Financial time series forecasting using independent component analysis and support vector regression," Decis. Support Syst., vol. 47, no. 2, pp. 115–125, 2009.

[28] M. Coli, R. Di Nisio, and L. Ippoliti, “Exploratory analysis of financial time series using independent component analysis,” in Information Technology Interfaces, 2005. 27th International Conference on, 2005, pp. 169–174.

[29] E. H. Wu and P. L. Yu, “Independent component analysis for clustering multivariate time series data,” 2005, pp. 474–482.

[30] S.-M. Cha and L.-W. Chan, “Applying independent component analysis to factor model in finance,” in IDEAL ’00: Proceedings of the Second International Conference on Intelligent Data Engineering and Automated Learning, Data Mining, Financial Engineering, and Intelligent Agents. Springer-Verlag, 2000, pp. 538–544.

[31] Kolenda, Independent components in text, ser. Advances in Independent Component Analysis. Springer-Verlag, 2000, pp. 229–250.

[32] E. Bingham, J. Kuusisto, and K. Lagus, “Ica and som in text document analysis,” in SIGIR ’02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2002, pp. 361–362.

[33] Q. Pu and G.-W. Yang, “Short-text classification based on ica and lsa,” Advances in Neural Networks-ISNN 2006, pp. 265–270, 2006.

[34] Kolenda, T., Hansen, L. K., and Sigurdsson, S. 2000. Independent components in text. In Advances in Independent Component Analysis, pp. 229–250. London: Springer-Verlag.

[35] Bingham, E., Kab´ an, A., and Girolami, M. 2001. Finding topics in dynamical text: application to chat line discussions. In Poster Proceedings of the 10th International World Wide Web Conference WWW10), pp. 198–199. Hong Kong: The Chinese University of Hong Kong.

[36] Timo Honkela, Aapo Hyvärinen, and Jaakko Väyrynen. WordICA - Emergence of linguistic representations for words by independent component analysis. Natural Language Engineering, 16(3):277–308, 2010.

[37] Hurri, J., Gäavert, H., Säareläa, J., and Hyväarinen, A. 2002. FastICA software package. Technical report, Laboratory of Computer and Information Science, Helsinki University of Technology, Espoo, Finland.

[38] Pearson, K. (1901). "On Lines and Planes of Closest Fit to Systems of Points in Space" . Philosophical Magazine 2 (11): 559–572.

[39] Sokal, R. R. and F. J. Rohlf. 1962. The comparison of dendrograms by objective methods. Taxon, 11:33-40

[40] Jain, A., and R. Dubes. Algorithms for Clustering Data. Upper Saddle River, NJ: Prentice-Hall, 1988.

[41] Zahn, C. T. "Graph-theoretical methods for detecting and describing Gestalt clusters." IEEE Transactions on Computers. Vol. C-20, Issue 1, 1971, pp. 68–86.

[42] http://www.behnevis.com

[43] Susan Schneider and Max Velmans (2008). "Introduction". In Max Velmans, Susan Schneider. The Blackwell Companion to Consciousness. Wiley. ISBN 978-0-470-75145-9.

[44] Mechanism (philosophy). (2014, May 25). In *Wikipedia, The Free Encyclopedia*. Retrieved 17:17, May 27, 2014, from http://en.wikipedia.org/w/index.php?title=Mechanism_(philosophy)&oldid=610089880

[45] Ned Block (1998). "On a confusion about a function of consciousness". In N. Block, O. Flanagan, G. Guzeldere. The Nature of Consciousness: Philosophical Debates. MIT Press. pp. 375–415. ISBN 978-0-262-52210-6.

[46] Consciousness. (2014, May 18). In *Wikipedia, The Free Encyclopedia*. Retrieved 19:09, May 27, 2014, from http://en.wikipedia.org/w/index.php?title=Consciousness&oldid=609088845

[47] David Chalmers (1995). "Facing up to the problem of consciousness". Journal of Consciousness Studies 2: 200–219.

[48] Fodor, J. A. (1980) Methodological solipsism considered as a research strategy in cognitive psychology. Behavioral and Brain Sciences 3: 63 - 109.

[49] Fodor, J. A. (1985) Précis of "The Modularity of Mind." Behavioral and Brain Sciences 8: 1 - 42.

[50] Newell, A. (1980) Physical Symbol Systems. Cognitive Science 4: 135 - 83.

[51] Pylyshyn, Z. W. (1984) Computation and cognition. Cambridge MA: MIT/Bradford

[52] Fodor, J. A. (1987) Psychosemantics Cambridge MA: MIT/Bradford.

[53] Kleene, S. C. (1969) Formalized recursive functionals and formalized realizability. Providence, R.: American Mathematical Society.

[54] Harnad, S. (1990) The Symbol Grounding Problem. Physica D 42: 335-346.

[55] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh and Mo Nours Arab (2008). “Human-Computer Interaction: Overview on State of the Art”. International Journal on Smart Sensing and Intelligent Systems, Vol. 1, No. 1, pp. 137-159.

[56] G. Robles-De-La-Torre, “The Importance of the sense of touch in virtual and real environments”, IEEE Multimedia 13(3), Special issue on Haptic User Interfaces for Multimedia Systems, pp 24-30 (2006).

[57] V. Hayward, O.R. Astley, M. Cruz-Hernandez, D. Grant and G. Robles-De-La-Torre, “Haptic interfaces and devices”, Sensor Review 24(1), pp 16-29 (2004).

[58] H. Iwata, “Haptic interfaces”, in J.A. Jacko and A. Sears (eds), The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Application, Lawrence Erlbaum Associates, Mahwah (2003).

[59] W. Barfield and T. Caudell, Fundamentals of Wearable Computers and Augmented Reality, Lawrence Erlbaum Associates, Mahwah (2001).

[60] W. Barfield and T. Caudell, Fundamentals of Wearable Computers and Augmented Reality, Lawrence Erlbaum Associates, Mahwah (2001).

[61] Global Positioning System, “Home page”, http://www.gps.gov/, visited on 31/5/2014.

[62] S.G. Burnay, T.L. Williams and C.H. Jones, Applications of Thermal Imaging, A. Hilger, Bristol (1988).

[63] J. Y. Chai, P. Hong and M. X. Zhou, “A probabilistic approach to reference resolution in multimodal user interfaces”, Proceedings of the 9th International Conference on Intelligent User Interfaces, Funchal, Madeira, Portugal, pp 70-77 (2004).

[64] S.L. Oviatt, P. Cohen, L. Wu, J. Vergo, L. Duncan, B. Suhm, J. Bers, T. Holzman, T. Winograd, J. Landay, J. Larson and D. Ferro, “Designing the user interface for multimodal speech and pen-based gesture applications: state-of-the-art systems and future research directions”, Human-Computer Interaction, 15, pp 263-322 (2000).

[65] M.N. Huhns and M.P. Singh (eds), Readings in Agents, Morgan Kaufmann, San Francisco (1998).

[66] C.S. Wasson, System Analysis, Design, and Development: Concepts, Principles, and Practices, John Wiley & Sons, Hoboken (2006).

[67] E.A. Bretz, “When work is fun and games”, IEEE Spectrum, 39(12), pp 50-50 (2002).

[68] J. Vince, Introduction to Virtual Reality, Springer, London (2004).

[69] K. McMenemy and S. Ferguson, A Hitchhiker’s Guide to Virtual Reality, A K Peters, Wellesley (2007).

[70] Xu, H., Chua, T.S.: Fusion of AV features and external information sources for event detection in team sports video. ACM Trans. Multimed. Comput. Commun. Appl. 2(1), 44–67 (2006)

[71] Vo, M.T. (1998). “A framework and Toolkit for the Construction of Multimodal Learning Interfaces”, PhD. Thesis, Carnegie Mellon University, Pittsburgh, USA.

[72] Pavlovic, V.I., Berry, G.A., Huang, T.S. (1997). “Integration of audio/visual information for use in human-computer intelligent interaction”. Proceedings of the 1997 International Conference on Image Processing (ICIP '97), Volume 1, pp. 121-124.

[73] Andre, M., Popescu, V.G., Shaikh, A., Medl, A., Marsic, I., Kulikowski, C., Flanagan J.L. (1998). “Integration of Speech and Gesture for Multimodal Human-Computer Interaction”. In Second International Conference on Cooperative Multimodal Communication. 28–30 January, Tilburg, The Netherlands.

[74] Cohen, P.R.; Johnston, M.; McGee, D.; Oviatt, S.L.; Pittman, J.; Smith, I.A.; Chen, L.; Clow, J. (1997). “Quickset: Multimodal interaction for distributed applications”, ACM Multimedia, pp. 31-40.

[75] Johnston, M.; Bangalore, S. (2000). “Finite-state Multimodal Parsing and Understanding”, In Proceedings of the International Conference on Computational Linguistics, Saarbruecken, Germany.

[76] Nigay, L.; Coutaz, J. (1995). “A generic platform for addressing the multimodal challenge”. Proceedings of the Conference on Human Factors in Computing Systems, ACM Press.

[77] Bouchet, J.; Nigay, L.; Ganille, T. (2004). “Icare software components for rapidly developing multimodal interfaces”. ICMI '04: Proceedings of the 6th international conference on Multimodal interfaces (New York, NY, USA), ACM, pp. 251-258.

[78] Russ, G., Sallans, B., Hareter, H. (2005). “Semantic Based Information Fusion in a Multimodal Interface”. International Conference on Human-Computer Interaction (HCI’05), Las Vegas, Nevada, USA, 20–23 June, pp 94-100.

[79] Vo, M.T., Wood, C. (1996). “Building an application framework for speech and pen input integration in multimodal learning interfaces”. In Proceedings of the Acoustics, Speech, and Signal Processing (ICASSP’96), May 7–10, IEEE Computer Society, Volume 06, pp. 3545-3548.

[80] Corradini, A., Mehta M., Bernsen, N.O., Martin, J.-C. (2003). “Multimodal Input Fusion in Human-Computer Interaction on the Example of the on-going NICE Project”. In Proceedings of the NATO-ASI conference on Data Fusion for Situation Monitoring, Incident Detection, Alert and Response Management, Yerevan, Armenia.

[81] Johnston, M.; Bangalore, S. (2000). “Finite-state Multimodal Parsing and Understanding”, In Proceedings of the International Conference on Computational Linguistics, Saarbruecken, Germany.

[82] Sun, Y., Shi, Y., Chen, F. and Chung , V.(2007). “An Efficient Multimodal Language Processor for Parallel Input Strings in Multimodal Input Fusion,” in Proc. of the international Conference on Semantic Computing, pp. 389-396.

[83] Reitter, D.; Panttaja, E. M.; Cummins, F. (2004). “UI on the fly: Generating a multimodal user interface,” in Proc. of HLT-NAACL-2004, Boston, Massachusetts, USA.

[84] D'Ulizia, A.; Ferri, F.; Grifoni P. (2011). “A Learning Algorithm for Multimodal Grammar Inference”, IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, Vol. 41 (6), pp. 1495 - 1510.

[85] Pérez, G.; Amores, G.; Manchón, P. (2005). “Two strategies for multimodal fusion”. In Proceedings of Multimodal Interaction for the Visualization and Exploration of Scientific Data, Trento, Italy, 26–32.

[86] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli. Multimodal fusion for multimedia analysis: A survey. Springer Multimedia Systems Journal, 16(6):345-379 (2010).

[87] Potamianos, G., Neti, C., Gravier, G., Garg, A., Senior, A.: Recent advances in the automatic recognition of audiovisual speech. Proc. IEEE 91(9), 1306–1326 (2003)

[88] Lucey, S., Sridharan, S., Chandran, V.: Improved speech recognition using adaptive audio-visual fusion via a stochastic secondary classifier. In: International Symposium on Intelligent Multimedia, Video and Speech Processing, pp. 551–554. Hong Kong (2001)

[89] Iyengar, G., Nock, H.J., Neti, C.: Audio-visual synchrony for detection of monologue in video archives. In: IEEE International Conference on Acoustics, Speech, and Signal Processing. Hong Kong (2003)

[90] Iyengar, G., Nock, H.J., Neti, C.: Discriminative model fusion for semantic concept detection and annotation in video. In: ACM International Conference on Multimedia, pp. 255–258. Berkeley (2003)

[91] Aleksic, P.S., Katsaggelos, A.K.: Audio-visual biometrics. Proc. IEEE 94(11), 2025–2044 (2006)

[92] Bredin, H., Chollet, G.: Audio-visual speech synchrony measure for talking-face identity verification. In: IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, pp. 233–236. Paris (2007)

[93] Neti, C., Maison, B., Senior, A., Iyengar, G., Cuetos, P., Basu, S., Verma, A.: Joint processing of audio and visual information for multimedia indexing and human-computer interaction. In: International Conference RIAO. Paris, France (2000)

[94] Jaffre, G., Pinquier, J.: Audio/video fusion: a preprocessing step for multimodal person identification. In: International Workshop on MultiModal User Authentification. Toulouse, France (2006)

[95] Wu, K., Lin, C.K., Chang, E., Smith, J.R.: Multimodal information fusion for video concept detection. In: IEEE International Conference on Image Processing, pp. 2391–2394. Singapore (2004)

[96] Guironnet, M., Pellerin, D., Rombaut, M.: Video classification based on low-level feature fusion model. In: The 13th European Signal Processing Conference. Antalya, Turkey (2005)

[97] Town, C.: Multi-sensory and multi-modal fusion for sentient computing. Int. J. Comput. Vis. 71, 235–53 (2007)

[98] Magalhaes, J., Ruger, S.: Information-theoretic semantic multimedia indexing. In: International Conference on Image and Video Retrieval, pp. 619–626. Amsterdam, The Netherlands (2007)

[99] Xu, H., Chua, T.S.: Fusion of AV features and external information sources for event detection in team sports video. ACM Trans. Multimed. Comput. Commun. Appl. 2(1), 44–67 (2006)

[100] Xie, L., Kennedy, L., Chang, S.F., Divakaran, A., Sun, H., Lin, C.Y.: Layered dynamic mixture model for pattern discovery in asynchronous multi-modal streams. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 2, pp. 1053–1056. Philadelphia, USA (2005)

[101] Yan, R., Yang, J., Hauptmann, A.: Learning query-class dependent weights in automatic video retrieval. In: ACM International Conference on Multimedia, pp. 548–555. New York, USA (2004)

[102] Shigeo Morishima and Satoshi Nakamura. 2004. Multimodal translation system using texture-mapped lip-sync images for video mail and automatic dubbing applications. EURASIP J. Appl. Signal Process. 2004 (January 2004), 1637-1647. http://dx.doi.org/10.1155/S1110865704404259

[103] Duygulu, P., Barnard, K., de Freitas, N., Forsyth, D., ”Object Recognition as Machine Translation: Learning a lexicon for a fixed image vocabulary”, Proc. European Conference on Computer Vision (ECCV2002), Copenhagen, Denmark, May 27 - June 2, 2002 (Lecture Notes in computer Science, Volume 2353/2002, pp.97).

[104] J. Jeon, V. Lavrenko and R. Manmatha, “Automatic Image Annotation and Retrieval Using Cross-media Relevance Models,” Proc. ACM SIGIR’03, pp 119-126, 2003.

[105] Zanasi, Alessandro (2009). "Virtual Weapons for Real Wars: Text Mining for National Security". Proceedings of the International Workshop on Computational Intelligence in Security for Information Systems CISIS'08. Advances in Soft Computing 53. p. 53. doi:10.1007/978-3-540-88181-0_7. ISBN 978-3-540-88180-3.

[106] Cohen, K. Bretonnel; Hunter, Lawrence (2008). "Getting Started in Text Mining". PLoS Computational Biology 4 (1): e20. doi:10.1371/journal.pcbi.0040020. PMC 2217579. PMID 18225946.

[107] Jenssen, Tor-Kristian; Lægreid, Astrid; Komorowski, Jan; Hovig, Eivind (2001). "A literature network of human genes for high-throughput analysis of gene expression". Nature Genetics 28 (1): 21–8. doi:10.1038/ng0501-21. PMID 11326270.

[108] Masys, Daniel R. (2001). "Linking microarray data to the literature". Nature Genetics 28 (1): 9–10. doi:10.1038/ng0501-9. PMID 11326264.

[109] Joseph, Thomas; Saipradeep, Vangala G; Venkat Raghavan, Ganesh Sekar; Srinivasan, Rajgopal; Rao, Aditya; Kotte, Sujatha; Sivadasan, Naveen (2012). "TPX: Biomedical literature search made easy". Bioinformation 8 (12): 578–80. doi:10.6026/97320630008578. PMC 3398782. PMID 22829734.

[110] Coussement, Kristof; Van Den Poel, Dirk (2008). "Integrating the voice of customers through call center emails into a decision support system for churn prediction". Information & Management 45 (3): 164–74. doi:10.1016/j.im.2008.01.005.

[111] Coussement, Kristof; Van Den Poel, Dirk (2008). "Improving customer complaint management by automatic email classification using linguistic style features as predictors". Decision Support Systems 44 (4): 870–82. doi:10.1016/j.dss.2007.10.010.

[112] Pang, Bo; Lee, Lillian; Vaithyanathan, Shivakumar (2002). "Thumbs up?". Proceedings of the ACL-02 conference on Empirical methods in natural language processing 10. pp. 79–86. doi:10.3115/1118693.1118704.

[113] Erik Cambria; Robert Speer, Catherine Havasi and Amir Hussain (2010). "SenticNet: a Publicly Available Semantic Resource for Opinion Mining". Proceedings of AAAI CSK. pp. 14–18.

[114] Cody, W.F., J.T. Kreulen, V. Krishna, W.S. Spangler. The integration of business intelligence and knowledge management. IBM SYSTEMS JOURNAL, Vol.41, No.4, 2002

[115] Courtney D. Corley, Diane J. Cook, Armin R. Mikler, Karan P. Singh. Text and Structural Data Mining of Influenza Mentions in Web and Social Media. Int. J. Environ. Res. Public Health 2010, 7, 596-615; doi:10.3390/ijerph7020596

[116] D. Zeng, et al., "Social Media Analytics and Intelligence," IEEE Intelligent Systems, vol. 25, no. 6, 2010, pp. 13–16.

[117] Dekang Lin. 1998. Automatic retrieval and clustering of similar words. In Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 2 (ACL '98), Vol. 2. Association for Computational Linguistics, Stroudsburg, PA, USA, 768-774. DOI=10.3115/980691.980696 http://dx.doi.org/10.3115/980691.980696

[118] Ido Dagan, Lillian Lee, and Fernando Pereira. 1997. Similarity-based method for word sense disambiguation. In Proceedings of the 35th Annual Meeting of the ACL, pages 56-63, Madrid, Spain.

[119] James R. Curran and Marc Moens. 2002. Improvements in automatic thesaurus extraction. In Proceedings of the ACL-02 workshop on Unsupervised lexical acquisition - Volume 9 (ULA '02), Vol. 9. Association for Computational Linguistics, Stroudsburg, PA, USA, 59-66. DOI=10.3115/1118627.1118635 http://dx.doi.org/10.3115/1118627.1118635

[120] http://www.lemurproject.org/

[121] A. McCallum. Bow: A toolkit for statistical language modeling, text retrieval, classification and clustering. http://www.cs.cmu.edu/~mccallum/bow, 1996.

[122] L. Baker, A. McCallum. Distributional Clustering of Words for Text Classification, ACM SIGIR Conference, 1998.

[123] R. Bekkerman, R. El-Yaniv, Y. Winter, N. Tishby. On Feature Distributional Clustering for Text Categorization. ACM SIGIR Conference, 2001.

[124] Charu C. Aggarwal, ChengXiang Zhai. (2012). "A Survey of Text Clustering Algorithms". In Mining Text Data (pp. 77-128). Springer US. http://dx.doi.org/10.1007/978-1-4614-3223-4_4

[125] Fodor, Jerry A. (1983). Modularity of Mind: An Essay on Faculty Psychology. Cambridge, Mass.: MIT Press. ISBN 0-262-56025-9

[126] Anil K. Jain and Richard C. Dubes. 1988. Algorithms for Clustering Data. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

[127] M. Sahlgren. The Word-Space Model: using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces. PhD thesis, Stockholm University, 2006.

[128] A.Balali, A. Rajabi, S. Ghasemi, M. Asadpour, H. Faili. "Content Diffusion Prediction in Social Networks" 5th International Conference on Information and Knowledge Technology, Iran, 2013

[129] Army of the Guardians of the Islamic Revolution. (2014, June 7). In Wikipedia, The Free Encyclopedia. Retrieved 19:11, June 9, 2014, from
http://en.wikipedia.org/w/index.php?title=Army_of_the_Guardians_of_the_Islamic_Revolution&oldid=611987344

[130] Husayn ibn Ali. (2014, June 9). In Wikipedia, The Free Encyclopedia. Retrieved 19:31, June 9, 2014, from http://en.wikipedia.org/w/index.php?title=Husayn_ibn_Ali&oldid=612197803

[131] P O Skjaervo. "Persian, Modern." Concise Encyclopedia of Languages of the World. Harvard University, Cambridge, MA, USA. 1st ed. 2009. Print

[132] P O Skjaervo. "Iranian Languages." Concise Encyclopedia of Languages of the World. Harvard University, Cambridge, MA, USA. 1st ed. 2009. Print

[133] Emmerick, Ronald Eric. “Iranian languages.” Encyclopedia Britannica Online. Retrieved July 5, 2014

[134] Afar, Janet. “Iran.” Encyclopedia Britannica Online. Retrieved July 5, 2014

[135] "Persian language." Encyclopædia Britannica. Encyclopædia Britannica Ultimate Reference Suite. Chicago: Encyclopædia Britannica, 2013.

[136] "Afghanistan." Encyclopædia Britannica. Encyclopædia Britannica Ultimate Reference Suite. Chicago: Encyclopædia Britannica, 2013.

[137] Persian vocabulary. (2014, April 24). In Wikipedia, The Free Encyclopedia. Retrieved 14:12, July 5, 2014, from http://en.wikipedia.org/w/index.php?title=Persian_vocabulary&oldid=605576894

[138] Scott Dyer, Jeff Martin, John Zulauf, “Motion Capture White Paper,” 12 Dec 1995, on-line at ftp://ftp.sgi.com/sgi/A%7CW/jam/mocap/MoCapWP_v2.0.html

[139] Maureen Furniss, “Motion Capture.” Retrieved from http://web.mit.edu/comm-forum/papers/furniss.html

[140] Motion capture. (2014, June 10). In Wikipedia, The Free Encyclopedia. Retrieved 12:32, June 20, 2014, from http://en.wikipedia.org/w/index.php?title=Motion_capture&oldid=612424957

[141] Förger, K. & Takala, T. & Pugliese, R. 2012. Authoring Rules for Bodily Interaction: From Example Clips to Continuous Motions. Conf. On Intelligent Virtual Agents, September 12-14, 2012, Santa Cruz, California, USA. LNCS, Intelligent Virtual Agents. Springer Berlin. 14 pages.

[142] Förger, K. & Honkela, T. & Takala, T. 2013. Impact of Varying Vocabularies on Controlling Motion of a Virtual Actor. Conf. On Intelligent Virtual Agents, August 29-31, 2013, Edinburgh, UK. LNCS, Intelligent Virtual Agents. Springer Berlin. 10 pages.

[143] John R. Searle (1980). Minds, brains, and programs. Behavioral and Brain Sciences, 3, pp 417-424. doi:10.1017/S0140525X00005756.

[144] Cole, David, "The Chinese Room Argument", *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2014/entries/chinese-room/>.

Appendix A. Adverbs combined

In this section, I have included the adverbs which have been combined.

Words	Meaning	Part of speech
لنگ_لنگان	limping	Adverb
لنگ_لنگ_زنان	limping	Adverb
کشان_کشان	limping	Adverb
پا_شلیدن	limping	Adverb
لنگان_لنگان	limping	Adverb
لنگان	limping	Adverb

The above table contains all the synonyms of "limping" adverb expressed by annotators. There have been also two rare cases in which two verbs have been used as an adverb. I have also considered that.

Words	Meaning	Part of speech
به آرامی	Slowly	Adverb
ارام	Slow	Adjective
کند	Slow	Adjective
یواش	Slowly	Adverb
به کندی	Slowly	Adverb
آهسته	Slow	Adjective
به آهستگی	Slowly	Adverb
نرم	Soft	Adjective
به نرمی	Softly	Adverb

The above table contains all the synonyms of the adverb "slowly". Unlike the "limping", annotators have used adjective or an adverb to describe the slowness of a motion.

Words	Meaning	Part of speech
عصبی	Nervous	Adjective
باعصبانیت	Angrily	Adverb
غضبناک	enraged	adjective

The above table contains all the synonyms of the adverb "angrily".

Words	Meaning	Part of speech
بی حال	Languid	Adjective
با بیحالی	Languidly	Adverb
با بی حالی	languidly	Adverb

The above table contains all the synonyms of the adverb "languidly".

Words	Meaning	Part of speech
سرخوش	Happy	Adjective
سر خوش	Happy	Adjective
با خوشحالی	Happily	Adverb
با شادی	Happily	Adverb
شاد	Happy	Adjective

The above table contains all the synonyms of the adverb “happily”.

Words	Meaning	Part of speech
با ناراحتی	Sadly	Adverb
ناراحت	sad	adjective
غمگین	Sad	Adjective
با اندوه	With grief, with sorrow	Adverb
پژمرده	Withered, faded	Adjective
با افسردگی	Depressingly	Adverb

The above table contains all the synonyms of the adverb “sadly”.

Words	Meaning	Part of speech
با شتاب	Hastily	Adverb
شتابان	Hastily	Adverb
تند	Speedily	Adverb
با سرعت	Fast	Adverb
سریع	Fast	Adjective
با عجله	Hurriedly	Adjective

The above table contains all the synonyms of the adverb “fast”.

Words	Meaning	Part of speech
طبیعی	Natural	Adverb
معمولی	Normal	Adverb

The above table contains all the synonyms of the adverb “normal”.

Appendix B. Frequency of verbs and adjectives in the corpora

There are two corpuses from which the word frequencies are extracted, the first large corpus is Alef dataset, and the second one is Tabnak.

In this part, we count the frequencies of verbs, and by verbs, I mean whatever in input to the first box of each page of the form. In English, the actions animated in the form can be simply described by one word; however, in Farsi, one might use a verb that contains two words or two morphemes separated using a zero-width-non-joiner. **Thus, the frequencies of substrings rather than words must be counted.**

In Farsi, one uses an infinitive to describe an action in one word; infinitives and their derivation word-forms are not equivalent in Farsi. Since an individual is asked to fill the forms using one word or expression and not a sentence, actions are described by infinitives rather than the word-forms. In addition, it is the word-forms instead of their corresponding infinitives which occur in a corpus.

People have utilized 12 different verbs in total to describe the motions they have seen in the forms. In order to count the frequencies of a verb, both its infinitive and all its inflections have been taken into account. In other words, when the inflection of an infinitive is observed, its frequency is increased by one.

Appendix C. Verb frequencies in both corpuses

In the below table, you can see the verbs frequencies in both Alef and Tabnak dataset.

Verb categories	verbs	Spell or grammatical error	Frequencies in form answers	Frequencies in the Alef corpus	Frequencies in the Tabnak corpus
Walk			1030	63	205
	راه رفتن (walk)	no	1021		
	قدم زدن (walk)	no	4		
	قدمم زدن (walk)	yes	1		
	راه رفتن با خوشحالی (walk happily)	no	1		
	راه رفت (he walked)	no	1		
	راه (path)	yes	1		
	را	yes	1		
Run			159	77	280
	دویدن (to run)	no	157		
	ورزش دویدن (to run+exercise as a noun)	yes	1		
	دویدن آهسته (walk slow)	yes	1		
limp			40	1	4
	لنگیدن (to limp)	no	29		
	لنگان لنگان رفتن (to limp)	no	9		
	لنگ زدن (to limp)	no	1		
	لنگیدن خسته (to limp + tired)	yes	1		
Think			4	717	2289

	فکر کردن (to think)	no	2		
	تصمیم گرفتن (to decide)	no	2		
To be purposeful			1	5	8
	هدفمند بودن (to be purposeful)	no	1		
To measure			2	4	11
	متراژ کردن (to measure)	yes	1		
	اندازه گرفتن (to measure)	no	1		
To drum			1	0	0
	طبل زدن در رژهٔ نظامی (to drum in a military parade)	no	1		
To exercise			2	19	89
	ورزش کردن (to exercise)	no	1		
	تمرین کردن (to practice)	no	1		
To lift			1	1	2
	وزنه کشیدن با پا (to lift using feet)	no	1		
To pose			1		2
	ژست گرفتن (to pose, to make gesture)	no	1		
To soar			1	0	1
	خیز گرفتن (to begin to rise, to begin to soar)	yes	1		
To carry			2	35	134
	حمل کردن (to carry)	no	1		
	حمل کردنک (to carry)	yes	1		

In the following sections, I have included the synonyms of the 10 most frequent verbs and verb-adverbs and their translation between Finnish, English, and Farsi language.

Appendix D. Automatic detection result synonyms

Verbs have been represented by a 124-dimensional data points. They can also be viewed as vectors with 124 components; the first component denotes the frequency of a verb in the first video; the other components are defined in the same way. In other words, videos act like the context for the verbs. Next, the distances among all verbs have been computed using cosine metric, and for each verb, 5 closest verbs have been extracted. They can represent synonyms; the number of relevant extracted synonyms can indicate how effective this method is.

D.I. Synonyms result of English annotations

Instead of choosing 0.3, the fixed neighborhood distance is set to 0.5. This number determines the radius of the neighborhood of a video; thus, a video has a larger neighborhood than the Farsi and Finnish cases.

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms
walking	392	-s	Walking	walk	Strolling	Ambling	wandering	5
		s	walk	Walking	Strolling	amble	wandering	5
		rs	Walking	Limping	walk	limping	scuffing	2
limping	202	-s	Limping	stagger	leaping	climbing	hobbling	4
		s	Limping	scuffing	stagger	limb	leaping	5
		rs	Limping	scuffing	walking	walk	moving	3
Limping	110	-s	limping	leaping	stagger	hobbling	scuffing	5
		s	limping	scuffing	moving	leaping	shuffling	5
		rs	limping	scuffing	walking	walk	Walking	2
running	74	-s	sprinting	Running	Sprinting	Rushing	Run	5
		s	Running	sprinting	jogging	run	Jogging	5
		rs	Running	Limping	jogging	scuffing	walking	2
jogging	68	-s	Jogging	lurch	a_morning_jog	swagger	start_running	3
		s	Jogging	running	lurch	a_morning_jog	Running	4
		rs	Jogging	Limping	running	scuffing	limping	2
Walking	51	-s	walking	stride	walk	striding	ramble	5
		s	walking	walk	striding	stride	strutting	5
		rs	walking	walk	Limping	limping	scuffing	2
scuffing	31	-s	Limping	limping	shuffle	edging	shuffling	4
		s	Limping	limping	shuffle	shuffling	edging	4
		rs	Limping	limping	walking	moving	shuffle	4
marching	24	-s	Marching	stamping	Storming_off	funny_walk	ample	3
		s	Marching	march	stamping	stomping	Storming_off	4
		rs	stomping	walking	Walking	Limping	walk	4
walk	21	-s	walking	Walking	striding	Strolling	Ambling	5
		s	walking	Walking	striding	Strolling	Ambling	5
		rs	walking	Walking	limping	Limping	scuffing	2

stomping	21	-s	Stamping	dancing	stamping	marching	Stomping	5
		s	Stamping	Stomping	Speed_Walk	lumber	angry_walk	4
		rs	marching	Limping	walking	limping	scuffing	2
indexical-grounding								44
pattern-grounding								46
randomized-grounding								25

In the above table, only the 10 most frequent verbs are demonstrated. For each one of these verbs, 5 closest verbs are extracted from verb-video matrix. If an extracted verb overlaps with meaning of the current verb, it is considered to be a good synonym. Number of good snonyms has been counted, and it turns out that in the indexical-grounding approach 44 semantically related verbs have been extracted. The maximum number of semantically related verbs that could have been extracted is 50; thus, using cosine distance measure on verb-video matrix has led to relatively good result. Furthermore, if pattern-grounding process is based on the original motion data, 2 more good synonyms are detected. Since the motion data was created by extracting suitable features from the motion videos, we can claim that grounding English verbs based on motion features makes sense. On the other hand, if the pattern-grounding is based on a random data, the performance of finding good synonyms deteriorates.

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms
walking_slowly	57	-s	walking_thoughtfully	walking_unpurposefully	walk_slowly	Walking_Joyfully	Ambling_Leisurely	2
		s	walking_very_slowly	walking_sadly	walking_carefully	Strolling_Slowly	walking_thoughtfully	2
		rs	Walking_Slowly	walking_sadly	walking_carefully	walking_very_slowly	Limping_Slowly	2
limping_slowly	28	-s	Hobbling_Slowly	limping_sadly	walking_asymmetrically	shuffling_wounded	scuffing_painstakingly	1
		s	limping_sadly	limping_painfully	scuffing_painstakingly	scuffing_slowly	Hobbling_Slowly	2
		rs	limping_painfully	walking_carefully	Limping_Slowly	walking_asymmetrically	scuffing_painstakingly	1
limping_painfully	21	-s	limping_slowly	Limping_Slowly	walk_injured	limping_sadly	Walking_Watchfully	0
		s	limping_slowly	Limping_Slowly	limping_sadly	limbing_very_slowly	Limping_Painfully	1
		rs	limping_slowly	Limping_Slowly	walking_carefully	walking_very_slowly	Limping_Normally	0
walking_very_slowly	20	-s	walking_thoughtfully	wandering_depressively	walk_thougthful	walking_walking_casually	walking_slightly_threateningly	0
		s	meandering_slowly	walking_slowly	Walking_Wondering	Strolling_Slowly	walking_idly	3
		rs	walking_carefully	walking_slowly	limping_painfully	limping_slowly	walking_sadly	1
lki ng s_ adl	18	-s	shuffling_mournfully	walking_unpurposefully	meandering_sadly	wandering_sadly	walking_with_doubt	2

		s	walking_slowly	walking_sorrowfully	walking_unpurposefully	amble_sadly	Limping_Very_slightly	2
		rs	walking_slowly	Walking_Slowly	walking_very_slowly	meandering_sadly	walking_carefully	1
walking_confidently	16	-s	walking_energetically	walking_quickly	walking_casual_walking	walking_briskly	walk_briskly	3
		s	walking_briskly	walking_normally	walking_energetically	Walking_Normally	walking_very_colorfully	2
		rs	walking_briskly	walking_slowly	walking_quickly	Walking_Slowly	walking_normally	1
Limping_Slowly	16	-s	Limping_Painfully	limping_slowly	walk_injured	limping_slow	Limping_Carefully	2
		s	limping_painfully	limping_slowly	limping_sadly	climbing_very_colorfully	limping_slow	2
		rs	limping_painfully	Limping_Painfully	limping_slowly	walking_carefully	walking_slowly	1
walking_normally	15	-s	Walking_Normally	walking_walk_leisurely	walking_vigorously	walking_stretched_legs	walking_saunter	2
		s	Walking_Normally	walking_walk_leisurely	walking_vigorously	walking_stretched_legs	walking_saunter	2
		rs	Walking_Normally	walking_quickly	walking_confidently	walking_briskly	walking_actively	1
walking_carefully	15	-s	walking_very_slowly	walking_slowly	limping_slowly	walking_leaping_from_left_leg	Walking_Carefully	1
		s	walking_slowly	walking_very_slowly	Limping_Normally	scuffing_slowly	walking_sadly	0
		rs	walking_very_slowly	walking_slowly	limping_slowly	limping_painfully	Limping_Normally	0
limping_fast	15	-s	leaping_fast	walking_without_crossing_feet	shuffling_active	scuffing_resolutely	Limping_strenuously	4
		s	limping_quickly	leaping_fast	limping_hurriedly	walking_impeded	Limping_Painfullly	3
		rs	limping_quickly	leaping_fast	limping_hurriedly	Limping_Quite_quickly	Limping_Painfullly	4
indexical-grounding								17
pattern-grounding								19
randomized-grounding								12

Pattern-grounding excels the other two approaches even when the combination of verbs and adverbs are analyzed. In addition to finding more good synonyms, the pattern-grounding approach also manges to make some of the potential synonyms closer to the analyzed word. For instance, if we focus on 'walking-very-slowly' which is the 4th most frequent verb-adverb, we would notice that the closest verb-adverb is also its best synonym.

D.II. Synonyms in Farsi annotations

Extracting similar verbs has been not been applied on Farsi verbs because the verb 'walking' is so frequent that it has occupied most the vector space. However, one can extract similar verb-adverbs from the verb_adverb-video matrix.

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms
راه_رفتن_آرام	389	-s	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_خسته	راه_رفتن_غمگین	راه_رفتن_بی_حوصله	1
		s	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_خسته	راه_رفتن_غمگین	راه_رفتن_خیلی_آرام	2
		rs	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_خسته	قدم_زدن_اهسته	راه_رفتن_لنگ_لنگان	2
راه_رفتن_لنگان_لنگان	346	-s	راه_رفتن_لنگ_لنگان	راه_رفتن_لنگان	لنگیدن_خسته	لنگیدن_آسیب_دیده	لنگان_لنگان_راه رفتن_خیلی_لنگیدن	3
		s	راه_رفتن_لنگ_لنگان	قدم_زدن_بالنگیدن	راه_رفتن_کمی_لنگان_لنگان	قدم_زدن_با_لنگیدن	لنگیدن_آسیب_دیده	4
		rs	راه_رفتن_لنگ_لنگان	راه_رفتن_کمی_تند	راه_رفتن_کمی_لنگان_لنگان	راه_رفتن_خسته	لنگیدن_خسته	2
راه_رفتن_کمی_تند	135	-s	راه_رفتن_تند	راه_رفتن_خوشحال	راه_رفتن_سریع	راه_رفتن_عصبانی	راه_رفتن_باعجله	3
		s	راه_رفتن_تند	راه_رفتن_خوشحال	راه_رفتن_سریع	راه_رفتن_باعجله	راه_رفتن_عصبانی	3
		rs	راه_رفتن_تند	راه_رفتن_لنگ_لنگان	راه_رفتن_لنگان_لنگان	راه_رفتن_سریع	راه_رفتن_کمی_لنگان_لنگان	2
راه_رفتن_ناراحت	117	-s	راه_رفتن_آرام	راه_رفتن_خسته	راه_رفتن_با_ناراحتی	راه_رفتن_غمگین	راه_رفتن_بی_حال	2
		s	راه_رفتن_با_ناراحتی	راه_رفتن_خسته	قدم_زدن_باخستگی	راه_رفتن_آرام	راه_رفتن_غمگین	2
		rs	راه_رفتن_آرام	راه_رفتن_خسته	راه_رفتن_آهسته	راه_رفتن_لنگ_لنگان	قدم_زدن_اهسته	0
راه_رفتن_خسته	84	-s	راه_رفتن_ناراحت	راه_رفتن_آرام	راه_رفتن_با_ناراحتی	راه_رفتن_بی_حال	راه_رفتن_غمگین	1
		s	راه_رفتن_ناراحت	قدم_زدن_اهسته	راه_رفتن_آرام	قدم_زدن_باخستگی	لنگان_لنگان_راه رفتن_خیلی_اهسته	1
		rs	راه_رفتن_آرام	راه_رفتن_ناراحت	راه_رفتن_آهسته	راه_رفتن_لنگ_لنگان	راه_رفتن_لنگان_لنگان	0
راه_رفتن_تند	79	-s	راه_رفتن_سریع	راه_رفتن_با_عجله	راه_رفتن_خوشحال	راه_رفتن_کمی_تند	راه_رفتن_با_سرعت	4
		s	راه_رفتن_سریع	راه_رفتن_با_عجله	راه_رفتن_خوشحال	راه_رفتن_کمی_تند	راه_رفتن_کمی_سریع	4

		rs	راه_رفتن_سریع	راه_رفتن_کمی_تند	راه_رفتن_با_عجله	راه_رفتن_خوشحال	راه_رفتن_لنگ_لنگان	3
راه_رفتن_کمی_لنگان_لنگان	76	-s	راه_رفتن_لنگ_لنگان	لنگیدن_آهسته	قدم_زدن_بالنگیدن	راه_رفتن_لنگان_لنگان	قدم_زدن_با_لنگیدن	5
		s	راه_رفتن_لنگان_لنگان	راه_رفتن_لنگ_لنگان	قدم_زدن_بالنگیدن	لنگیدن_آسیب_دیده	لنگیدن_آهسته	4
		rs	راه_رفتن_لنگ_لنگان	لنگیدن_آهسته	راه_رفتن_لنگان_لنگان	راه_رفتن_آرام	لنگیدن_آرام	4
راه_رفتن_معمولی	72	-s	راه_رفتن_عادی	راهرفتن_به_طوری_نرمال	راه_رفتن_با_خوشحالی	راه_رفتن_نرمال	راه_رفتن_جدی	3
		s	راه_رفتن_عادی	راه_رفتن_با_خوشحالی	راهرفتن_به_طوری_نرمال	راه_رفتن_نرمال	راه_رفتن_نرم	4
		rs	راه_رفتن_عادی	راهرفتن_به_طوری_نرمال	راه_رفتن_آرام	راه_رفتن_با_خوشحالی	راه_رفتن_تند	2
راه_رفتن_با_عصبانیت	65	-s	راه_رفتن_بچه_گانه	راه_رفتن_موزون	گام_برداشتن_عصبی	راه_رفتن_مستانه	راه_رفتن_کوبنده	1
		s	راه_رفتن_بچه_گانه	راه_رفتن_موزون	گام_برداشتن_عصبی	راه_رفتن_مستانه	قدم_برداشتن_عصبی	2
		rs	راه_رفتن_بچه_گانه	راه_رفتن_عصبانی	راه_رفتن_محکم	راه_رفتن_موزون	راه_رفتن_باعصبانیت	1
راه_رفتن_آهسته	65	-s	راه_رفتن_آرام	راه_رفتن_غمگین	راه_رفتن_ناراحت	راه_رفتن_به_آرامی	راه_رفتن_با_ناراحتی	2
		s	راه_رفتن_آرام	راه_رفتن_بافکر	راه_رفتن_بی_حوصله	راه_رفتن_به_آرامی	راه_رفتن_افسرده	2
		rs	راه_رفتن_آرام	راه_رفتن_خسته	راه_رفتن_ناراحت	قدم_زدن_اهسته	راه_رفتن_افسرده	2
indexical-grounding								25
pattern-grounding								28
randomized-grounding								18

More relevant synonyms are detected when pattern-grounding is implemented. Furthermore, in general, the correct synonyms would get closer to the studied verb-adverb by pattern-grounding; for example, 'راه_رفتن_با_ناراحتی' is ranked the 3rd closest verb-adverb to 'راه_رفتن_ناراحت', while with the help of pattern-grounding, it is ranked the 1st.

D.III. Synonyms in Finnish annotation

The synonyms of both Finnish verbs and verb-adverbs have been extracted. In the following table, you can see the synonyms of the 10 most frequent Finnish verbs and verb-adverbs.

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms

kävelee	618	-s	kävellä	käppäilee	käveleminen	Kävelee	löntystelee	5
		s	kävellä	käppäilee	Kävelee	löntystelee	maleksii	5
		rs	kävellä	nilkuttaa	ontua	ontuu	linkuttaa	1
ontuu	229	-s	nilkuttaa	ontua	linkuttaa	laahustaa	raahustaa	3
		s	ontua	nilkuttaa	linkuttaa	raahustaa	laahustaa	3
		rs	nilkuttaa	ontua	linkuttaa	laahustaa	kävelee	3
nilkuttaa	164	-s	ontuu	linkuttaa	ontua	linkkaa	liikkuu	4
		s	ontuu	linkuttaa	ontua	raahustaa	laahaa_jalkaa	4
		rs	ontuu	ontua	linkuttaa	kävelee	kävellä	3
juoksee	106	-s	juosta	juokseminen	ryntää	pyrähtää	lähtee_juoksemaan	5
		s	juosta	juokseminen	lähtee_juoksemaan	ryntää	pyrähtää	5
		rs	juosta	juokseminen	kävelee	ontua	ontuu	2
kävellä	82	-s	kävelee	käppäilee	käveleminen	reippailee	löntystelee	5
		s	kävelee	käppäilee	maleksii	astelee	käyskentelee	5
		rs	kävelee	nilkuttaa	ontuu	ontua	linkuttaa	1
hölkkää	66	-s	hölkätä	hölkyttää	hölkkääminen	Hölkkää	jolkottelee	4
		s	hölkyttää	hölkätä	lönkyttelee	jolkottelee	hölkkääminen	4
		rs	hölkyttää	hölkätä	kävelee	juosta	kävellä	3
ontua	52	-s	ontuu	nilkuttaa	linkuttaa	laahustaa	raahustaa	3
		s	ontuu	raahustaa	nilkuttaa	laahustaa	linkuttaa	3
		rs	ontuu	nilkuttaa	linkuttaa	kävelee	laahustaa	3
linkuttaa	45	-s	ontuu	nilkuttaa	ontua	liikkuu	linkkaa	4
		s	nilkuttaa	ontuu	ontua	raahustaa	liikkuu	4
		rs	ontuu	nilkuttaa	ontua	laahustaa	raahustaa	3
laahustaa	34	-s	ontuu	klenkkaa	raahustaa	ontua	raahautuu	2
		s	ontua	raahustaa	ontuu	klenkkaa	raahautuu	2
		rs	ontuu	nilkuttaa	ontua	linkuttaa	raahustaa	1
juosta	16	-s	juoksee	juokseminen	ryntää	pyrähtää	ottaa_spurtit	5
		s	juoksee	juokseminen	lähtee_juoksemaan	ryntää	pyrähtää	5
		rs	juoksee	kävelee	ontua	nilkuttaa	ontuu	1
indexical-grounding								40
pattern-grounding								40
randomized-grounding								21

Although the resulting synonyms using indexical-grounding is not precisely identical to the results of pattern-grounding approach, there is no profound difference between these two approaches in terms of number of found relevant synonyms. However, the performance of pattern-grounding approach would deteriorate significantly when Finnish verbs are grounded using a random data.

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good synonyms
kävelee_hitaasti	64	-s	kävellä_hitaasti	kävelee_surullisena	kävelee_epävarmasti	kävelee_hyvin_hitaasti	käveleskelee_mietteliäästi	2
		s	kävellä_hidas	kävelee_hyvin_hitaasti	kävellä_hitaasti	kävelee_surullisena	maleksii_hitaasti	4
		rs	kävellä_hitaasti	kävelee_surullisena	kävelee_mietteliäästi	kävelee_varovasti	kävelee_rennosti	1
kävelee_reippaasti	52	-s	kävellä_rivakasti	kävelee_määrätietoisesti	kävelee_päättäväinen	kävelee_kävelee_normaalisti	Kävelee_Reippaasti	4
		s	kävelee_määrätietoisesti	kävelee_päättäväinen	kävellä_rivakasti	kävellä_reippaasti	kävellä_rento	4
		rs	kävelee_määrätietoisesti	kävelee_tavallisesti	kävelee_päättäväinen	kävellä_reipas	kävellä_rivakasti	4
ontuu_hitaasti	47	-s	nilkuttaa_hitaasti	ontuu_varovasti	ontuu_vaivalloisesti	nilkuttaa_kivulloisesti	linkuttaa_vaivainen	1
		s	nilkuttaa_hitaasti	ontuu_varovasti	ontuu_hyvin_kivuliaasti	ontuu_vaivalloisesti	ontuu_takavetoisesti	1
		rs	nilkuttaa_hitaasti	ontuu_varovasti	kävelee_vaivalloisesti	ontuu_vaivalloisesti	ontuu_pahasti	1
kävelee_ontuen	41	-s	nilkuttaa_nopeasti	kävelee_reippahasti	ontuu_reippaasti	ontuu_takavetoisesti	nilkuttaa_reippaasti	3
		s	ontuu_pahasti	kävelee_rammiten	nilkuttaa_nopeasti	laahaa_jalkaa_pahasti	kävelee_laahustaen	4
		rs	kävelee_rauhallisesti	kävelee_rammiten	kävelee_vaivalloisesti	ontuu_hitaasti	kävelee_hitaasti	1
kävelee_rennosti	40	-s	kävellä_itsevarmasti	kävelee_itsevarmasti	kävelee_jäykästi	kävelee_kiireettä	kävellä_rennosti	1
		s	kävelee_itsevarmasti	kävellä_itsevarmasti	kävellä_rennosti	kävelee_verkkaisesti	vetelehtii_hitaasti	3
		rs	kävelee_hitaasti	kävellä_rennosti	kävelee_itsevarmasti	kävelee_normaalisti	kävelee_rauhallisesti	2
kävelee_normaalisti	28	-s	kävellä_itsevarma	kävellä_normaali	kävellä_itsevarmasti	kävelee_tavallisesti	kävellä_rento	2
		s	kävellä_itsevarma	kävellä_rento	kävelee_tavallisesti	kävelee_jäykästi	kävelee_kävelee_normaalisti	2
		rs	kävelee_rennosti	kävellä_itsevarma	kävelee_tavallisesti	kävelee_reippaasti	kävelee_määrätietoisesti	1
ontuu_vaivalloisesti	21	-s	linkuttaa_vaivainen	ontuu_hitaasti	nilkuttaa_hitaasti	ontuu_hyvin_kivuliaasti	ontuu_kivualiaasti	5
		s	linkuttaa_vaivainen	nilkuttaa_hitaasti	ontuu_hyvin_kivuliaasti	ontuu_hitaasti	ontua_surullinen	4
		rs	ontuu_hitaasti	nilkuttaa_hitaasti	ontuu_hyvin_kivuliaasti	ontuu_kivuliaasti	nilkuttaa_vaikeasti	5
[illegible]	20	-s	ontuu_vaike	linkuttaa_kiv	ontuu_voima	ontuu_vihais	Ontuu_vaiva	3

			asti	ulias	kkaasti	esti	lloisesti	
		s	ontuu_vaikeasti	nilkuttaa_vaikeasti	linkuttaa_kivulias	raahustaa_vaikeasti	ontuu_vaivaloisesti	5
		rs	ontuu_vaikeasti	nilkuttaa_hitaasti	ontuu_hitaasti	ontuu_vaivaloisesti	kävelee_vaivalloisesti	4
kävelee_rauhallisesti	15	-s	kävelee_kiireettömästi	kävelee_hieman_alakuloisesti	kävellä_epävarmasti	kävelee_tyynesti	vaeltaa_hitaasti	3
		s	kävelee_hitaasti	kävelee_kiireettömästi	kävellä_hidas	kävelee_hieman_alakuloisesti	kävellä_epävarmasti	3
		rs	kävelee_hitaasti	kävelee_ontuen	kävelee_rammiten	kävelee_rennosti	ontuu_hitaasti	2
juoksee_rennosti	15	-s	juoksee_kohtuullista_vauhtia	juosta_rennosti	juoksee_reippaasti	juoksee_juoksee_hitaasti	hölkkää_rennosti	4
		s	hölkkää_reippaasti	juoksee_kohtuullista_vauhtia	juoksee_rauhallisesti	juosta_rennosti	juoksee_reippaasti	4
		rs	juoksee_reippaasti	juosta_rennosti	ontuu_hitaasti	hölkkää_hitaasti	kävelee_varovasti	1
indexical-grounding								28
pattern-grounding								34
randomized-grounding								22

When the combination of Finnish verb and adverbs are considered, more relevant synonyms are extracted using the pattern-grounding approach. Even for less frequent verb-adverbs such as ‘juoksee_rennosti’, the pattern-grounding approach assigns a lower rank to inappropriate synonyms such as ‘juoksee_reippaasti’.

Appendix E. Automatic translation result

In Appendix, you can find the translation result of the 10 most frequent verbs and verb-adverbs. The translation has been implemented using indexical-grounding, pattern-grounding, and randomized-grounding approach. In addition, English, Finnish, and Farsi annotations have been been considered in the translation process.

Verb translation has been implemented by considering a verb vector in the source language and the closest verb vectors in the target language. Five of such closest verbs have been reported in the following tables. Verb-adverbs have been translated using the same process.

E.I. Translating from English to Finnish

In this section, the result of translating from English to Finnish is declared.

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	#good translation
walking	392	-s	kävelee	kävellä	Kävelee	käveleminen	käppäilee	5
		s	kävelee	kävellä	Kävelee	käppäilee	astelee	5

		rs	kävelee	kävellä	ontuu	nilkuttaa	ontua	2
limping	202	-s	ontuu	nilkuttaa	linkuttaa	ontua	liikkuu	5
		s	nilkuttaa	ontuu	linkuttaa	ontua	raahustaa	4
		rs	ontuu	nilkuttaa	linkuttaa	ontua	kävelee	4
Limping	110	-s	ontuu	nilkuttaa	linkuttaa	ontua	liikkuu	5
		s	ontuu	nilkuttaa	ontua	linkuttaa	raahustaa	4
		rs	ontuu	nilkuttaa	ontua	linkuttaa	kävelee	4
running	74	-s	juoksee	juosta	juokseminen	ryntää	pyrähtää	5
		s	juoksee	juosta	juokseminen	starttaa_juoksuun	hölkkää	5
		rs	juoksee	juosta	kävelee	nilkuttaa	hölkkää	3
jogging	68	-s	hölkkää	hölkätä	hölkyttää	hölkkääminen	Hölkkää	5
		s	hölkkää	hölkyttää	hölkätä	lönkyttelee	jolkottaa	5
		rs	hölkkää	hölkyttää	juosta	nilkuttaa	linkuttaa	3
Walking	51	-s	kävellä	kävelee	reippailee	Kävelee	käveleminen	4
		s	kävellä	kävelee	astelee	käppäilee	Kävelee	5
		rs	kävellä	kävelee	ontuu	nilkuttaa	ontua	2
scuffing	31	-s	ontuu	ontua	nilkuttaa	raahustaa	linkuttaa	1
		s	ontua	ontuu	raahustaa	laahustaa	nilkuttaa	2
		rs	ontuu	nilkuttaa	ontua	linkuttaa	kävelee	0
marching	24	-s	marssii	tramppaaminen	marssia	harppoo	tömistää	4
		s	marssii	tömistää	harppoo	tramppaaminen	marssia	4
		rs	harppoo	kävelee	ontuu	kävellä	linkuttaa	3
walk	21	-s	kävellä	kävelee	Kävelee	käveleminen	hidastelee	4
		s	kävelee	kävellä	käppäilee	Kävelee	astelee	5
		rs	kävellä	kävelee	nilkuttaa	ontuu	linkuttaa	2
stomping	21	-s	tömistelee	tömpsyttelee	tömpsii	tömistely	tömistellä	5
		s	tömistelee	polkee_jalkaa	tömpsyttelee	tömpsii	tömistely	5
		rs	kävelee	ontuu	nilkuttaa	linkuttaa	ontua	1
indexical-grounding								43
pattern-grounding								44
randomized-grounding								24

Pattern-grounding approach has led to slightly better translation from English verb to Finnish.

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
ng_sl	57	-s	kävelee_hitaasti	kävellä_hitaasti	löntystelee_rennosti	käveleskelee_mietteliääs	kävelee_rennosti	4

					ti		
		s	kävelee_hitaasti	kävellä_hidas	kävellä_hitaasti	kävelee_rauhallisesti	kävelee_hyvin_hitaasti	5
		rs	kävelee_hitaasti	kävellä_hitaasti	kävelee_rennosti	kävelee_rauhallisesti	ontuu_hitaasti	4
limping_slowly	28	-s	ontuu_hitaasti	nilkuttaa_hitaasti	ontuu_raskaasti	ontuu_alakuloisesti	nilkuttaa_kivulloisesti	2
		s	ontuu_hitaasti	nilkuttaa_hitaasti	ontuu_varovasti	ontuu_vaivalloisesti	laahustaa_ontuen	3
		rs	ontuu_hitaasti	ontuu_varovasti	ontuu_vaivalloisesti	nilkuttaa_hitaasti	ontua_hitaasti	3
limping_painfully	21	-s	nilkuttaa_kivulloisesti	laahustaa_vaivalloisesti	ontuu_hitaasti	ontuu_väsyneesti	raahustaa_tuskaisesti	3
		s	nilkuttaa_hitaasti	ontuu_hitaasti	ontuu_vaivalloisesti	ontuu_hyvin_kivuliaasti	ontua_surullinen	2
		rs	ontuu_hitaasti	ontuu_vaivalloisesti	kävelee_rammiten	ontuu_pahasti	ontuu_hyvin_kivuliaasti	2
walking_very_slowly	20	-s	kävelee_hitaasti	kävelee_epävarmasti	kävelee_surullisena	tallustaa_miettien	pohdiskella_mietteliäs	1
		s	kävelee_hyvin_hitaasti	kävelee_hitaasti	kävelee_epävarmasti	käyskentelee_mietteliäästi	kävelee_varovasti	2
		rs	kävelee_hitaasti	kävelee_surullisena	ontuu_hitaasti	kävelee_epävarmasti	kävellä_hitaasti	2
walking_sadly	18	-s	kävelee_allapäin	matelee_allapäin	Kävelee_Hitaasti	kävelee_alakuloisesti	kävellä_pohdiskeleva	1
		s	kävellä_pohdiskeleva	kävelee_allapäin	kävelee_mietteliäästi	kävelee_surullisena	kävellä_hitaasti	1
		rs	kävelee_mietteliäästi	kävelee_masentuneesti	kävelee_surullisena	kävelee_hitaasti	kävellä_hitaasti	2
walking_confidently	16	-s	kävellä_reippaasti	kävelee_päättäväinen	kävelee_reippaasti	kävelee_iloisesti	kävellä_rivakasti	1
		s	kävelee_normaalisti	kävellä_rento	kävellä_reippaasti	kävelee_määrätietoisesti	kävelee_reippaasti	1
		rs	kävelee_päättäväinen	kävellä_reippaasti	kävelee_reippaasti	kävelee_määrätietoisesti	kävelee_tavallisesti	2
Limping_Slowly	16	-s	nilkuttaa_varoen	ontuu_hitaasti	nilkuttaa_hitaasti	ontuu_kivuliaasti	kävelee_raahautuen	2
		s	ontuu_hitaasti	nilkuttaa_hitaasti	ontua_hitaasti	ontuu_varovasti	ontuu_raskaasti	3
		rs	ontuu_hitaasti	ontuu_vaivalloisesti	ontuu_hyvin_kivuliaasti	ontuu_kivuliaasti	ontuu_varovasti	1
walking_normally	15	-s	reippailee_tarmokkaasti	nilkuttaa_hyvin_hyvin_vähän	kävellä_pystypäin	käveleminen_jäykästi	käveleminen_epäluonnollisesti	0
		s	reippailee_tarmokkaasti	nilkuttaa_hyvin_hyvin_vähän	kävellä_pystypäin	käveleminen_jäykästi	käveleminen_epäluonnollisesti	0

		rs	kävelee_normaalisti	kävellä_reippaasti	kävellä_normaali	kävellä_itsevarmasti	kävelee_rennosti	3
walking_carefully	15	-s	kävelee_hitaasti	kävelee_epävarmasti	ontua_varovainen	kävelee_varovasti	ontuu_hitaasti	1
		s	kävelee_hitaasti	kävelee_epävarmasti	kävelee_väsyneesti	kävelee_kävelee_nilkuttaen_hitaasti	kävelee_varovasti	1
		rs	kävelee_hitaasti	ontuu_hitaasti	kävelee_rennosti	kävelee_varovasti	kävelee_ontuen	1
limping_fast	15	-s	ontuu_vauhdikkaasti	ontuu_päämäärätietoisesti	ontuu_kiireisesti	ontuu_askel_kerrallaan	ontua_huoleton	2
		s	ontuu_reippaasti	ontuu_ripeästi	nilkuttaa_reippaasti	nilkuttaa_nopeasti	ontua_reipas	5
		rs	nilkuttaa_nopeasti	ontuu_reippaasti	nilkuttaa_reippaasti	kävelee_ontuen	kävelee_linkuttaen	3
indexical-grounding								17
pattern-grounding								23
randomized-grounding								23

Pattern-grounding enhances the automatic ranslation of English verb-adverb to Finnish verb-adverb. However, since a randomized-grounding also has led to a comparable translation performance, it cannot be said with certain that it was pattern-grounding that enhances the translation in this case.

E.II. Translating from Finnish to English

In this section, you ca find the translation result of 10 most frequent annotated Finnish verbs and verb-adverbs to English.

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
kävelee	618	-s	walking	Walking	walk	striding	Ambling	5
		s	walking	Walking	walk	Strolling	amble	5
		rs	walking	Walking	Limping	walk	limping	3
ontuu	229	-s	limping	Limping	scuffing	stagger	leaping	3
		s	limping	Limping	scuffing	shuffling	limb	4
		rs	limping	Limping	scuffing	walking	moving	3
nilkuttaa	164	-s	limping	Limping	stagger	scuffing	leaping	3
		s	limping	Limping	limbing	scuffing	limb	4
		rs	Limping	limping	scuffing	walking	moving	3
juoksee	106	-s	running	sprinting	Running	Sprinting	Rushing	5
		s	running	Running	sprinting	Starting_running	Sprinting	5
		rs	running	Running	sprinting	jogging	walking	4
kävell	82	-s	walking	Walking	walk	Ambling	striding	5

ä		s	walking	Walking	walk	Strolling	amble	5
		rs	walking	Walking	walk	Limping	limping	3
hölkkä ä	66	-s	jogging	Jogging	lurch	a_morning_j og	swagger	3
		s	jogging	Jogging	lurch	a_morning_j og	trot	4
		rs	jogging	Jogging	running	run	Limping	4
ontua	52	-s	limping	Limping	scuffing	stagger	leaping	3
		s	Limping	limping	scuffing	shuffling	Hobbling	4
		rs	Limping	limping	scuffing	walking	moving	3
linkutt aa	45	-s	limping	Limping	limp	shuffling	climbing	4
		s	limping	Limping	scuffing	limbing	leaping	3
		rs	limping	Limping	scuffing	walking	moving	3
laahus taa	34	-s	shuffling	trodding	limping	Limping	scuffing	2
		s	scuffing	shuffling	edging	Limping	plodding	3
		rs	scuffing	Limping	limping	walking	moving	2
juosta	16	-s	running	Running	sprinting	Sprinting	Rushing	5
		s	running	Running	sprinting	Starting_run ning	Sprinting	5
		rs	running	Running	jogging	sprinting	walking	4
indexical-grounding								38
pattern-grounding								42
randomized-grounding								32

Using pattern-grounding, more appropriate translations have been found, while using a randomized-grounding, the translation performance is worsened in comparison to the indexical-grounding approach.

Verb-adverb	Frequency	Mode	1st extracted verb-adv	2nd	3rd	4th	5th	# good translation
kävelee_hitaa sti	64	-s	walking_slo wly	walking_tho ughtfully	walking_unp urposefully	walking_ver y_slowly	walking_wai ting	3
		s	walking_slo wly	walking_ver y_slowly	walking_tho ughtfully	walking_idly	walking_wai ting	4
		rs	walking_slo wly	walking_ver y_slowly	walking_car efully	Walking_Slo wly	limping_pai nfully	3
kävelee_reipp aasti	52	-s	Walking_Pu rposefully	walking_ere ct	walking_ene rgetically	walking_fast	walking_con fidently	2
		s	walking_stea dily	walking_bris kly	Walking_Pu rposefully	walking_ere ct	walking_acti vely	3
		rs	walking_con fidently	walking_fast	walking_bris kly	walking_stea dily	walking_ene rgetically	4
ontuu_hitaa sti	47	-s	limping_slo wly	limping_sadl y	Hobbling_Sl owly	shuffling_w ounded	limping_pai nfully	2
		s	limping_slo wly	limping_pai nfully	limping_sadl y	scuffing_slo wly	scuffing_pai nstakingly	1
		rs	limping_slo	limping_pai	Limping_Slo	scuffing_slo	walking_slo	2

			wly	nfully	wly	wly	wly	
kävelee_ontuen	41	-s	walking_limpingly	walking_impeded	Limping_Quickly	walking_uncoordinnated	walking_slightly_limply	3
		s	limping_hurriedly	limping_customarily	walking_impeded	Limping_Very_fast	Limping_Slowly	1
		rs	Limping_Very_fast	limping_hurriedly	Limping_Normally	Limping_Slowly	limping_painfully	1
kävelee_rennosti	40	-s	walking_regularly	walking_slowly	walking_normally	Walking_Normally	Striding_Determined	3
		s	walking_relaxed	walking_very_slow	walking_cautious	strolling_lazily	moving_on_foot	3
		rs	walking_slowly	walking_briskly	Walking_Slowly	walking_carefully	Limping_Slowly	0
kävelee_normaalisti	28	-s	walking_normal	walking_normally	walk_purposefully	strutting_confident	walking_regularly	3
		s	walking_normal	walking_confidently	walking_regularly	walking_normally	walking_briskly	3
		rs	walking_normal	walking_normally	Walking_Quickly	walking_briskly	walking_confidently	2
ontuu_vaivalloisesti	21	-s	Limping_Painfully	limping_slowly	limping_with_difficulty	walking_leaping_from_left_leg	walk_difficulty	2
		s	limping_painfully	limping_slowly	Limping_Very_badly	limping_with_difficulty	limb_slowly	3
		rs	limping_slowly	Limping_Slowly	limping_painfully	limping_slow	Limping_Painfully	2
ontuu_kivuliaasti	20	-s	limping_with_difficulty	walk_injured	walk_difficulty	limping_difficultly	limb_strongly	2
		s	limping_with_difficulty	Limping_Badly	walk_difficulty	Limping_Very_badly	limping_heavily	3
		rs	Limping_Slowly	limping_painfully	limping_slowly	walk_injured	limping_slow	1
kävelee_rauhallisesti	15	-s	walking_slowly	walking_relaxedly	wandering_indecisively	walk_thoughtfully	walking_walking_comfortably	2
		s	walking_slowly	strolling_slowly	walking_relaxedly	Walking_Slowly	loafing_carelessly	2
		rs	walking_slowly	walking_relaxed	Walking_Slowly	Limping_Slowly	walking_briskly	1
juoksee_rennosti	15	-s	running_very_colorfully	running_running	running_fast	swagger_cool	start_running_slowly	1
		s	running_very_colorfully	running_fast	running_powerfully	Jogging_Normally	Running_Quite_slowly	1
		rs	running_very_colorfully	running_running	jogging_slowly	running_fast	Jogging_Lightly	1
indexical-grounding								23
pattern-grounding								24
randomized-grounding								17

In addition to finding one more appropriate translation, pattern-grounding slightly boosts the translation of Finnish verb-adverbs by assigning the more accurate translations higher ranks. For example, ‘walking_regularly’ is ranked the 5th best translation of ‘kävelee_normaalisti’ using the indexical-grounding approach, while pattern-grounding improves its rank and raise it to the 3rd best translation.

E.III. Translating from English to Farsi

In this section, you can find the translation of English verbs and verb-adverbs to Farsi verbs and verb-adverbs.

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
walking	392	-s	راه_رفتن	راهرفتن	قدم_زدن	تند_راه_رفتن	راه_می_رود	5
		s	راه_رفتن	راهرفتن	قدم_زدن	پیاده_روی	گام_برداشتن	5
		rs	راه_رفتن	قدم_زدن	گام_برداشتن	راهرفتن	لنگیدن	4
limping	202	-s	لنگیدن	لنگان_لنگان_رفتن	لنگان_لنگان_راه رفتن	لنگان_لنگان_راه_رفتن	می_لنگد	5
		s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_رفتن	قدم_برداشتن	لنگان_لنگان_راه_رفتن	4
		rs	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	قدم_زدن	گام_برداشتن	2
Limping	110	-s	لنگیدن	لنگان_لنگان_رفتن	لنگان_لنگان_راه رفتن	می_لنگد	لنگان_لنگان_راه_رفتن	5
		s	لنگان_لنگان_راه رفتن	لنگیدن	لنگان_لنگان_رفتن	لنگان_لنگان_راه_رفتن	راه_رفتن	5
		rs	لنگیدن	راه_رفتن	لنگان_لنگان_راه رفتن	قدم_زدن	گام_برداشتن	3
running	74	-s	دویدن	می_دود	سریع_دویدن	دویدن_سریع	با_سرعت_دویدن	5
		s	دویدن	می_دود	سریع_دویدن	دویدن_سریع	با_سرعت_دویدن	5
		rs	دویدن	می_دود	سریع_دویدن	راه_رفتن	قدم_زدن	3
jogging	68	-s	دویدن_آهسته	آهسته_دویدن	دویدن	نرم_دویدن	آهسته_دمیدن	5
		s	دویدن_آهسته	دویدن	آهسته_دویدن	نرم_دویدن	آهسته_دمیدن	5
		rs	دویدن	هروله	دویدن_آهسته	آهسته_دویدن	قدم_زدن	4
Walking	51	-s	تند_راه_رفتن	راهرفتن	راه_رفتن	پیاده_روی	راه_رفتن_تیز	5
		s	راهرفتن	راه_رفتن	تند_راه_رفتن	پیاده_روی	قدم_زدن	5
		rs	راه_رفتن	قدم_زدن	راهرفتن	گام_برداشتن	لنگیدن	4
scuffing	31	-s	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	lang_langan_rah_raftan	لنگان_لنگان_راه_رفتن	0
		s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_رفتن	کشال_کشال_راه_رفتن	لنگان_لنگان_راه_رفتن	1
		rs	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	قدم_زدن	گام_برداشتن	0

marching	24	-s	هدفمند_بودن	راه_رفتن_بصورت_چکشی	خبره	تیز_و_قوی_گشتن	نرمش_کردن	0
		s	نرمش_کردن	رژه_رفتن	هدفمند_بودن	راه_رفتن_بصورت_چکشی	خبره	1
		rs	گام_برداشتن	راه_رفتن	راهرفتن	قدم_زدن	رژه_رفتن	5
walk	21	-s	راه_رفتن	راهرفتن	قدم_زدن	تند_راه_رفتن	لنگان_لنگان_رفتن	4
		s	راه_رفتن	راهرفتن	قدم_زدن	تند_راه_رفتن	گام_برداشتن	5
		rs	راه_رفتن	قدم_زدن	راهرفتن	لنگیدن	گام_برداشتن	4
stomping	21	-s	عصبانی_راه_رفتن	طبل_زدن_در_رژه_نظامی	سریع_راه_رفتن	رقصیدن	راه_رفتن_با_عصبانیت	1
		s	عصبانی_راه_رفتن	طبل_زدن_در_رژه_نظامی	سریع_راه_رفتن	رقصیدن	راه_رفتن_با_عصبانیت	1
		rs	گام_برداشتن	راه_رفتن	راهرفتن	قدم_زدن	لنگیدن	0
indexical-grounding								35
pattern-grounding								37
randomized-grounding								29

Translation of English verbs and verb-adverbs to Farsi has been enhanced moderately when pattern-grounding is applied. Since randomized-grounding has worsened the translation, it can be claimed that pattern-grounding using the motion data is a meaningful and reasonable way of normalization.

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
walking_slowly	57	-s	راه_رفتن_آرام	قدم_زدن_آهسته	راه_رفتن_آهسته	قدم_زدن_آرام	راه_رفتن_با_آرامش	4
		s	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_بی_حوصله	راه_رفتن_اهسته	قدم_زدن_آهسته	4
		rs	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_بافکر	راه_رفتن_ناراحت	راه_رفتن_بی_حوصله	2
limping_slowly	28	-s	لنگان_لنگان_رفتن_بد	لنگان_لنگان_رفتن_آرام	لنگان_لنگان_رفتن_شل	لنگان_لنگان_رفتن_خراب	راه_رفتن_پیر	2
		s	لنگیدن_آهسته	لنگیدن_خسته	لنگیدن_خیلی_آهسته	لنگیدن_ارام	لنگیدن_آرام	4
		rs	لنگیدن_خسته	راه_رفتن_خسته	لنگیدن_آهسته	راه_رفتن_لنگان_لنگان	لنگان_لنگان_رفتن_لنگان	3
limping_painfully	21	-s	لنگان_لنگان_رفتن_بد	لنگان_لنگان_رفتن_شل	لنگان_لنگان_رفتن_خراب	لنگیدن_پادرد	لنگان_لنگان_رفتن_لنگان	3
		s	لنگیدن_خسته	قدم_زدن_بالنگی	راه_رفتن_لنگان	قدم_برداشتن_اها	راه_رفتن_لنگ_	3

				دن	_لنگان	سته	لنگان	
		rs	قدم_زدن_بالنگیدن	راه_رفتن_لنگ_لنگان	راه_رفتن_لنگان_لنگان	راه_رفتن_خسته	راه_رفتن_آرام	3
walking_very_slowly	20	-s	راه_رفتن_متفکرانه	راه_رفتن_آرام	راه_رفتن_بی_حوصله	راه_رفتن_متفکرانه	راه_رفتن_خسته	1
		s	راه_رفتن_متفکرانه	راه_رفتن_آهسته	راه_رفتن_متفکرانه	قدم_زدن_با_ناراحتی	راه_رفتن_آرام	2
		rs	راه_رفتن_متفکرانه	راه_رفتن_خسته	راه_رفتن_آرام	راه_رفتن_ناراحت	راه_رفتن_آهسته	2
walking_sadly	18	-s	راه_رفتن_شکست_خورده	قدم_زدن_بیحال	سرگردان_فقیر	رفتن_بی_خیل_رفتن	راهرفتن_ناراحتی	2
		s	راه_رفتن_افسرده	راه_رفتن_غمگین	راه_رفتن_به_آرامی	راه_رفتن_شکست_خورده	راه_رفتن_ناراحت	4
		rs	راه_رفتن_بی_حال	راه_رفتن_با_ناراحتی	راه_رفتن_به_آرامی	راه_رفتن_ناراحت	راه_رفتن_خسته	2
walking_confidently	16	-s	راه_رفتن_بیخیال	راه_رفتن_نرم	راه_رفتن_با_اعتماد_به_نفس	راه_رفتن_نرمال	راه_رفتن_سبک	1
		s	راهرفتن_به_طوری_نرمال	راه_رفتن_با_خوشحالی	راه_رفتن_معمولی	راه_رفتن_عادی	راه_رفتن_نرمال	0
		rs	راه_رفتن_تند	راه_رفتن_سریع	راه_رفتن_با_اعتماد_به_نفس	راه_رفتن_با_خوشحالی	راهرفتن_به_طوری_نرمال	1
Limping_Slowly	16	-s	لنگان_لنگان_رفتن_لنگان	لنگان_لنگان_رفتن_آرام	راه_رفتن_لنگان_لنگان	راه_رفتن_با_مصدومیت	لنگان_لنگان_رفتن_کند	4
		s	راه_رفتن_لنگان_لنگان	راه_رفتن_لنگ_لنگان	لنگان_لنگان_رفتن_لنگان	قدم_زدن_بالنگیدن	لنگیدن_اهسته	5
		rs	لنگان_لنگان_رفتن_لنگان	راه_رفتن_لنگان_لنگان	راه_رفتن_لنگ_لنگان	لنگیدن_خسته	راه_رفتن_آرام	4
walking_normally	15	-s	گام_بلند_برداشتن_گامهای_بلند_برداشتن	گام_بلند_برداشتن_عادی_گام_برداشتن	گام_برداشتن_عجول	رژه_رفتن_منظم	رژه_رفتن_تند_تند	2
		s	گام_بلند_برداشتن_گامهای_بلند_برداشتن	گام_بلند_برداشتن_عادی_گام_برداشتن	گام_برداشتن_عجول	رژه_رفتن_منظم	رژه_رفتن_تند_تند	2
		rs	راهرفتن_به_طوری_نرمال	راه_رفتن_معمولی	راه_رفتن_با_سرعت	راه_رفتن_سبک	راه_رفتن_با_خوشحالی	2
walking_carefully	15	-s	راه_رفتن_آرام	راه_رفتن_آهسته	لنگان_لنگان_راهرفتن_با_عصبانیت	راه_رفتن_به_سختی	راه_رفتن_ناراحت	0
		s	راه_رفتن_آرام	قدم_زدن_باخس	راه_رفتن_ناراح	راه_رفتن_آهسته	راه_رفتن_اهسته	0

				تگی	ت			
		rs	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_خسته	راه_رفتن_لنگان_لنگان_	0
limping_fast	15	-s	لنگیدن_عجله	لنگان_لنگان_سریعرفتن_خراب	راه_رفتن_نفس_نفس_زنان	راه_رفتن_تند_راه_رفتن	می_لنگد_لنگلنگان_راه_می_رود	3
		s	لنگیدن_عجله	لنگان_لنگان_سریعرفتن_خراب	راه_رفتن_نفس_نفس_زنان	راه_رفتن_تند_راه_رفتن	لنگیدن_سریع	3
		rs	لنگیدن_عجله	لنگان_لنگان_سریعرفتن_خراب	راه_رفتن_نفس_نفس_زنان	راه_رفتن_تند_راه_رفتن	راه_رفتن_لنگیدن	3
indexical-grounding								22
pattern-grounding								27
randomized-grounding								22

In addition to finding more appropriate translations, using pattern-grounding approach, the correct translations have been ranked higher in comparison to the tranlsations found by the other two approaches.

E.IV. Translting from Farsi to Enlish

In this section, you can find the translation result from Farsi to English for the 10 most frequent verbs and verb-adverbs.

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
راه_رفتن	1637	-s	walking	walk	Walking	Limping	limping	3
		s	walking	walk	Walking	Limping	moving	4
		rs	walking	Limping	limping	Walking	scuffing	2
دویدن	339	-s	running	jogging	Running	Jogging	run	5
		s	running	jogging	Running	Jogging	sprinting	5
		rs	running	jogging	Running	walking	Jogging	4
لنگیدن	312	-s	limping	Limping	scuffing	stagger	hobbling	4
		s	limping	Limping	scuffing	limbing	limb	4
		rs	Limping	limping	scuffing	walking	moving	3
قدم_زدن	140	-s	walking	Strolling	wandering	falter	walk	4
		s	walking	Strolling	strolling	loiter	walk	5
		rs	walking	Limping	walk	limping	Walking	3
لنگان_لنگان_راهرفتن	36	-s	limping	Limping	moving	limb	scuffing	4
		s	Limping	limping	scuffing	moving	shuffling	4
		rs	Limping	limping	scuffing	walking	moving	3
گام_برداشتن	29	-s	marching	walking	Walking	march	stomping	5
		s	marching	walking	walk	march	stomping	5

		rs	walking	Limping	limping	Walking	scuffing	2
لنگان_لنگان_رفتن	23	-s	limping	Limping	stagger	leaping	hobbling	4
		s	limping	Limping	stagger	leaping	hobbling	4
		rs	limping	stagger	hobbling	Limping	leaping	4
راهرفتن	18	-s	walking	Walking	walk	Strolling	Ambling	5
		s	walking	Walking	walk	striding	Strolling	5
		rs	walking	Walking	walk	Limping	limping	3
قدم_برداشتن	11	-s	limbing	stompping	Start_walking	scuffing	Limping	2
		s	limb	limping	limbing	Limping	hobble	0
		rs	Limping	limping	limbing	walking	scuffing	1
هروله	9	-s	trot	prancing	jogging	moving	Limping	3
		s	trot	prancing	jogging	Jogging	moving	4
		rs	Limping	walking	moving	limbing	limping	1
indexical-grounding								39
pattern-grounding								40
randomized-grounding								26

When pattern-grounding is exploited, the translation of Farsi verbs into English is slightly enhanced, while randomized-grounding worsen the translation.

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
راه_رفتن_آرام	389	-s	walking_slowly	walking_very_slowly	walking_sadly	walking_carefully	walking_thoughtfully	2
		s	walking_slowly	walking_carefully	walking_very_slowly	Limping_Normally	walking_sadly	2
		rs	walking_slowly	limping_painfully	Limping_Slowly	Limping_Very_fast	Walking_Slowly	2
راه_رفتن_لنگان_لنگان	346	-s	Limping_Slowly	limping_slowly	Limping_Painfully	walk_injured	limping_painfully	0
		s	limping_painfully	Limping_Slowly	limping_slowly	Limping_Painfully	scuffing_very_slowly	0
		rs	limping_painfully	Limping_Slowly	Limping_Very_fast	walking_slowly	limping_slowly	0
راه_رفتن_کمی_تند	135	-s	Limping_Very_fast	limping_hurriedly	walking_actively	walking_unsteady	walking_slightly_weirdly	1
		s	Limping_Very_fast	Striding_Fast	walking_briskly	walking_heavily	Walking_Quickly	3
		rs	Limping_Very_fast	Limping_Normally	walking_briskly	limping_painfully	limping_hurriedly	1
راه_رفتن_ناراحت	117	-s	walking_sadly	walking_very_slowly	walk_depressed	walking_slowly	walking_sorrowfully	3
		s	walking_carefully	Limping_Normally	walking_slowly	walking_sadly	scuffing_slowly	1
		rs	limping_pai	walking_slo	Limping_No	Limping_Ve	Walking_Sl	0

			nfully	wly	rmally	ry_fast	owly	
راه_رفتن_خسته	84	-s	limping_slowly	walking_very_slowly	walking_sadly	Hobbling_Slowly	scuffing_slowly	0
		s	Limping_Normally	scuffing_slowly	walking_carefully	walking_very_slowly	walking_slowly	0
		rs	limping_painfully	Limping_Slowly	walking_slowly	Limping_Normally	scuffing_slowly	0
راه_رفتن_تند	79	-s	walking_confidently	walking_fast	walking_energetically	walk_briskly	Walking_Sedately	2
		s	walking_briskly	walking_confidently	walking_actively	walking_steadily	Walking_Quickly	3
		rs	walking_briskly	Limping_Very_fast	walking_confidently	Limping_Normally	Walking_Slowly	1
راه_رفتن_کمی_لنگان_لنگان	76	-s	walking_very_carefully	Limping_Slightly	walking_slightly_impeded	Limping_Normally	Limping_Very_fast	2
		s	Limping_Slowly	limping_painfully	Limping_Normally	Limping_Slightly	limping_slowly	1
		rs	Limping_Normally	Limping_Very_fast	limping_painfully	limping_hurriedly	scuffing_slowly	0
راه_رفتن_معمولی	72	-s	walking_normally	Walking_Normally	walking_briskly	walking_confidently	walking_energetically	2
		s	walking_normally	walking_confidently	walking_briskly	walking_normal	Walking_Normally	3
		rs	walking_briskly	walking_normally	Limping_Very_fast	limping_painfully	walking_slowly	1
راه_رفتن_با_عصبانیت	65	-s	marching_angrily	walking_angry	stomping_annoyed	walking_angrily	walking_weirdly	4
		s	marching_angrily	walking_angry	stomping_annoyed	walking_aggressively	Stamping_Angrily	5
		rs	marching_angrily	walking_angrily	Stamping_Angrily	walking_angry	stomping_annoyed	5
راه_رفتن_آهسته	65	-s	walking_slowly	walking_very_slowly	Walking_Slowly	walking_thoughtfully	walking_carefully	3
		s	walking_slowly	walking_very_slowly	walking_carefully	walking_reserved	walking_sadly	3
		rs	walking_slowly	Walking_Slowly	limping_painfully	Limping_Normally	Limping_Slowly	1
indexical-grounding								19
pattern-grounding								21
randomized-grounding								11

Pattern-grounding slightly engances the translation of Farsi verb-adverbs into English verb-adverbs, while randomized-grounding deteriorates the quality of translation.

E.V. Translating Farsi to Finnish

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation

راه_رفتن	1637	-s	kävelee	kävellä	nilkuttaa	ontuu	ontua	2
		s	kävelee	kävellä	löntystelee	Kävelee	nilkuttaa	4
		rs	kävelee	kävellä	nilkuttaa	ontuu	ontua	2
دویدن	339	-s	juosta	juoksee	hölkkää	juokseminen	Juoksee	5
		s	juosta	juoksee	hölkkää	starttaa_juoksuun	juokseminen	5
		rs	juosta	juoksee	hölkkää	hölkyttää	nilkuttaa	4
لنگیدن	312	-s	ontuu	nilkuttaa	ontua	linkuttaa	raahustaa	4
		s	ontuu	nilkuttaa	ontua	linkuttaa	raahustaa	4
		rs	nilkuttaa	ontuu	ontua	kävelee	linkuttaa	4
قدم_زدن	140	-s	kävelee	kävellä	hidastelee	maleksii	löntystelee	5
		s	kävelee	maleksii	kävellä	löntystelee	käyskentelee	5
		rs	kävelee	kävellä	nilkuttaa	ontuu	ontua	2
لنگان_لنگان_راهرفتن	36	-s	ontuu	nilkuttaa	ontua	linkuttaa	laahustaa	4
		s	ontuu	ontua	nilkuttaa	raahustaa	linkuttaa	4
		rs	ontuu	nilkuttaa	ontua	kävelee	linkuttaa	4
گام_برداشتن	29	-s	kävelee	harppoo	kävellä	marssii	nilkuttaa	4
		s	harppoo	kävelee	marssii	tömistää	kävellä	5
		rs	kävelee	kävellä	nilkuttaa	ontua	ontuu	2
لنگان_لنگان_رفتن	23	-s	ontuu	liikkuu	linkuttaa	nilkuttaa	ontua	5
		s	ontuu	linkuttaa	ontua	nilkuttaa	raahustaa	4
		rs	linkuttaa	ontuu	ontua	konkkaa	raahustaa	3
راهرفتن	18	-s	kävelee	kävellä	Kävelee	käppäilee	löntystelee	5
		s	kävellä	kävelee	käppäilee	Kävelee	harppoo	5
		rs	kävellä	kävelee	ontua	nilkuttaa	ontuu	2
قدم_برداشتن	11	-s	polkee_jalkaa	laahaa_jalkaa	haastaa_riitaa	raahautuu	tömistää	2
		s	nilkuttaa	ontuu	linkuttaa	ontua	raahustaa	0
		rs	nilkuttaa	ontua	kävelee	linkuttaa	ontuu	1
هروله	9	-s	kiirehtiä	hölkyttää	hölkkää	lönköttelee	hölkyttelee	3
		s	kiirehtiä	hölkkää	hölkyttää	hölkyttelee	hypähtelee	4
		rs	kävelee	nilkuttaa	kävellä	juosta	ontuu	1
indexical-grounding								39
pattern-grounding								40
randomized-grounding								25

The translation performance has been enhanced slightly by applying pattern-grounding, while it has been profoundly deteriorated using randomized-grounding.

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
آرام_رفتن	389	-s	kävelee_hitaasti	kävellä_hitaasti	kävelee_miettelíäästi	kävelee_varovasti	kävelee_surullisena	2

		s	kävelee_hitaasti	kävelee_epävarmasti	kävelee_kävelee_nilkuttaen_hitaasti	kävelee_rauhallisesti	kävelee_hyvin_hitaasti	3
		rs	kävelee_hitaasti	ontuu_hitaasti	kävelee_rammiten	kävelee_ontuen	kävelee_rennosti	1
راه_رفتن_لنگان_لنگان	346	-s	kävelee_ontuen	ontuu_kivuliaasti	ontuu_hitaasti	ontuu_vaivalloisesti	nilkuttaa_hitaasti	1
		s	ontuu_pahasti	kävelee_ontuen	ontuu_hitaasti	ontuu_kivuliaasti	kävelee_rammiten	1
		rs	kävelee_ontuen	ontuu_hitaasti	ontuu_hyvin_kivuliaasti	ontuu_pahasti	kävelee_rammiten	1
راه_رفتن_کمی_تند	135	-s	kävellä_reipas	kävelee_tomerasti	kävelee_ontuen	kävelee_hiukan_ontuen	kävelee_reippaasti	3
		s	kävelee_reippaasti	kävelee_tomerasti	kävelee_nopeasti	kävelee_tarmokkaasti	kävellä_reippaasti	5
		rs	kävelee_ontuen	ontuu_pahasti	kävellä_reipas	kävelee_rammiten	kävelee_reippaasti	2
راه_رفتن_ناراحت	117	-s	kävelee_mietteliäästi	kävellä_surullinen	kävelee_masentuneesti	kävelee_surullisena	kävelee_epävarmasti	3
		s	kävelee_epävarmasti	kävelee_kävelee_nilkuttaen_hitaasti	kävelee_väsyneesti	kävelee_masentuneesti	kävellä_surullinen	2
		rs	ontuu_hitaasti	kävelee_hitaasti	kävelee_rammiten	ontuu_pahasti	kävelee_ontuen	0
راه_رفتن_خسته	84	-s	ontuu_hitaasti	nilkuttaa_hitaasti	ontua_hitaasti	ontuu_varovasti	ontuu_raskaasti	0
		s	kävelee_kävelee_nilkuttaen_hitaasti	kävelee_epävarmasti	nilkuttaa_hitaasti	ontuu_hyvin_kivuliaasti	ontua_hitaasti	0
		rs	ontuu_hitaasti	kävelee_ontuen	ontuu_pahasti	ontuu_hyvin_kivuliaasti	kävelee_rammiten	0
راه_رفتن_تند	79	-s	kävelee_reippaasti	kävelee_päättäväinen	kävellä_reippaasti	kävelee_urheilullisesti	kävelee_iloisesti	2
		s	kävelee_reippaasti	kävelee_päättäväinen	kävelee_määrätietoisesti	kävellä_rento	kävellä_reippaasti	2
		rs	kävelee_reippaasti	kävelee_ontuen	kävellä_rento	kävellä_reipas	ontuu_pahasti	2
راه_رفتن_کمی_لنگان_لنگان	76	-s	kävelee_nilkuttaen	kävelee_ontuen	ontua_toispuoleinen	ontuu_pahasti	kävelee_laahustaen	3
		s	kävelee_ontuen	ontuu_varovasti	ontuu_pahasti	ontuu_hitaasti	ontua_hidas	1
		rs	kävelee_ontuen	ontuu_pahasti	kävelee_rammiten	kävelee_kävelee_nilkuttaen_hitaasti	ontuu_hyvin_kivuliaasti	2
راه_رفتن_معمولی	72	-s	kävelee_normaalisti	kävellä_normaali	kävelee_rennosti	kävellä_rento	kävelee_jäykästi	4
		s	kävelee_normaalisti	kävellä_rento	kävellä_itsevarma	kävelee_rennosti	kävelee_tavallisesti	4
		rs	kävellä_rento	kävelee_normaalisti	kävellä_itsevarma	kävelee_rennosti	kävelee_ontuen	3
	65	-s	kävelee_vih	kävelee_käv	kävelee_raiv	kävellä_kiuk	kävelee_kiu	4

			aisesti	elee_epänormaalisti	okkaasti	utellen	kkuisesti	
		s	kävelee_raivokkaasti	kävelee_vihaisesti	kävelee_kävelee_epänormaalisti	kävellä_kiukutellen	kävelee_kiukkuisesti	4
		rs	kävelee_kävelee_epänormaalisti	kävelee_vihaisesti	kävellä_ripeästi	kävelee_raivokkaasti	kävelee_tarmokkaasti	2
راه_رفتن_آهسته	65	-s	kävelee_hitaasti	kävelee_surullisena	kävelee_epävarmasti	kävellä_hitaasti	kävelee_varovasti	2
		s	kävelee_hitaasti	kävelee_epävarmasti	kävellä_hidas	kävellä_hitaasti	kävelee_rauhallisesti	4
		rs	kävelee_hitaasti	kävelee_ontuen	ontuu_hitaasti	kävelee_rennosti	kävelee_varovasti	2
indexical-grounding								24
pattern-grounding								26
randomized-grounding								15

Pattern-grounding has slightly bossted the translation of Farsi verb-adverbs to Finnish verb-adverbs, while randomized-grounding has strongly deteriorated the translation.

E.VI. Translating Finnish annotations to Farsi

Verb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	#good translation
kävelee	618	-s	راه_رفتن	راهرفتن	قدم_زدن	گام_برداشتن	تند_راه_رفتن	4
		s	راه_رفتن	راهرفتن	قدم_زدن	گام_برداشتن	لنگان_لنگان_راه رفتن	4
		rs	راه_رفتن	قدم_زدن	گام_برداشتن	راهرفتن	لنگان_لنگان_راه رفتن	4
ontuu	229	-s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_رفتن	راه_رفتن	لنگان_لنگان_راه_رفتن	5
		s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_رفتن	کشال_کشال_راه_رفتن	قدم_برداشتن	5
		rs	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	قدم_زدن	گام_برداشتن	5
nilkuttaa	164	-s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_رفتن	راه_رفتن	می_لنگد	5
		s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_رفتن	قدم_برداشتن	لنگان_لنگان_راه_رفتن	5
		rs	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	قدم_زدن	گام_برداشتن	5
juoksee	106	-s	دویدن	می_دود	سریع_دویدن	دویدن_سریع	با_سرعت_دویدن	5

		s	دویدن	می_دود	سریع_دویدن	دویدن_سریع	با_سرعت_دویدن	5
		rs	دویدن	می_دود	سریع_دویدن	راه_رفتن	قدم_زدن	3
		-s	راهرفتن	راه_رفتن	قدم_زدن	تند_راه_رفتن	راه_می_رود	5
kävellä	82	s	راهرفتن	راه_رفتن	قدم_زدن	پیاده_روی	آهسته_راه_رفتن	5
		rs	راه_رفتن	قدم_زدن	راهرفتن	گام_برداشتن	لنگان_لنگان_راه رفتن	4
		-s	دویدن_آهسته	آهسته_دویدن	دویدن	نرم_دویدن	آهسته_دمیدن	5
hölkkää	66	s	دویدن_آهسته	دویدن	آهسته_دویدن	نرم_دویدن	آهسته_دمیدن	5
		rs	دویدن	هروله	دویدن_آهسته	آهسته_دویدن	قدم_زدن	4
		-s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_رفتن	راه_رفتن	پایش_درد_می_کند	4
ontua	52	s	لنگیدن	لنگان_لنگان_راه رفتن	لنگان_لنگان_رفتن	کشال_کشال_راه_رفتن	پایش_درد_می_کند	4
		rs	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	قدم_زدن	گام_برداشتن	5
		-s	لنگیدن	لنگان_لنگان_رفتن	لنگان_لنگان_راه_رفتن	لنگان_لنگان_راه رفتن	شلیدن	5
linkuttaa	45	s	لنگیدن	لنگان_لنگان_رفتن	لنگان_لنگان_راه رفتن	لنگان_لنگان_راه_رفتن	قدم_برداشتن	5
		rs	لنگیدن	لنگان_لنگان_راه رفتن	راه_رفتن	قدم_زدن	گام_برداشتن	5
		-s	لنگان_لنگان_راه رفتن	گام_برداشتن_آهسته	لنگان_لنگان_رفتن	لنگیدن	راه_رفتن	1
laahustaa	34	s	لنگان_لنگان_راه رفتن	لنگیدن	پایش_درد_می_کند	لنگیدن_شدید	لنگان_لنگان_رفتن	0
		rs	لنگان_لنگان_راه رفتن	لنگیدن	قدم_زدن	راه_رفتن	گام_برداشتن	1
		-s	دویدن	می_دود	سریع_دویدن	ورزش_دویدن	دویدن_سریع	5
juosta	16	s	دویدن	می_دود	سریع_دویدن	دویدن_سریع	با_سرعت_دویدن	5
		rs	دویدن	می_دود	سریع_دویدن	راه_رفتن	قدم_زدن	3
indexical-grounding								44
pattern-grounding								43
randomized-grounding								39

Except 'laahustaa', other Finnish verbs have been translated quite well into Farsi; one reason for this good translation is that Finnish verbs are specific expression of some motion, while the Farsi verbs which have been used to annotate these motions are more general expression.

Verb-adverb	Frequency	Mode	1st extracted verb	2nd	3rd	4th	5th	# good translation
kävelee_hitaasti	64	-s	راه_رفتن_متفکرانه	قدم_زدن_آهسته	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_بی_حوصله	3
		s	راه_رفتن_متفکرانه	راه_رفتن_بی_حوصله	قدم_زدن_آهسته	راه_رفتن_آهسته	راه_رفتن_بی_هدف	2
		rs	راه_رفتن_متفکرانه	راه_رفتن_آرام	راه_رفتن_آهسته	راه_رفتن_ناراحت	راه_رفتن_بی_حوصله	2
kävelee_reippaasti	52	-s	راه_رفتن_مصمم	راه_رفتن_تند	راه_رفتن_سریع	راه_رفتن_مغرور	راهرفتن_به_طوری_نرمال	2
		s	راه_رفتن_مصمم	راه_رفتن_خوشحال	راه_رفتن_با_عجله	راه_رفتن_با_انرژی	راه_رفتن_تند	3
		rs	راه_رفتن_سریع	راه_رفتن_تند	راه_رفتن_با_عجله	راه_رفتن_مصمم	راه_رفتن_خوشحال	3
ontuu_hitaasti	47	-s	لنگیدن_خسته	لنگان_لنگان_رفتن_بد	لنگیدن_آهسته	راه_رفتن_لنگان_لنگان	لنگان_لنگان_رفتن_آرام	3
		s	لنگیدن_آهسته	لنگیدن_خسته	لنگیدن_آرام	لنگیدن_خیلی_آهسته	لنگیدن_به_آرامی	4
		rs	لنگیدن_خسته	لنگیدن_آهسته	راه_رفتن_لنگان_لنگان	راه_رفتن_خسته	راه_رفتن_آرام	2
kävelee_ontuen	41	-s	راه_رفتن_لنگان_لنگان	لنگان_لنگان_راهرفتن_با_تندی	راه_رفتن_لنگان	لنگیدن_سریع	لنگیدن_معلول	2
		s	راه_رفتن_لنگان_لنگان	راه_رفتن_لنگان	راه_رفتن_لنگ_لنگان	راه_رفتن_کمی_لنگان_لنگان	قدم_زدن_بالنگیدن	5
		rs	راه_رفتن_لنگان_لنگان	راه_رفتن_کمی_تند	راه_رفتن_لنگان	راه_رفتن_لنگ_لنگان	لنگیدن_سریع	3
kävelee_rennosti	40	-s	قدم_زدن_با_آرامی	راه_رفتن_نرمال	راه_رفتن_با_غرور	راه_رفتن_معمولی	راه_رفتن_عادی	4
		s	قدم_زدن_متفکرانه	راه_رفتن_با_آرامش	راه_رفتن_عادی	قدم_زدن_با_آرامی	راه_رفتن_معمولی	4
		rs	راه_رفتن_معمولی	راه_رفتن_آرام	راه_رفتن_عادی	راه_رفتن_آهسته	راه_رفتن_با_خوشحالی	4
kävelee_normaalisti	28	-s	راه_رفتن_معمولی	راه_رفتن_عادی	راه_رفتن_با_خوشحالی	راهرفتن_به_طوری_نرمال	راه_رفتن_ملایم	4
		s	راه_رفتن_معمولی	راه_رفتن_عادی	راه_رفتن_نرمال	راه_رفتن_ملایم	راهرفتن_به_طوری_نرمال	5
		rs	راه_رفتن_با_خو	راه_رفتن_معمو	راهرفتن_به_طور	راه_رفتن_عادی	راه_رفتن_تند	3

			شحالی	لی	ی_نرمال			
ontuu_vaivalloisesti	21	-s	لنگان_لنگان_رفتن_آرام	لنگیدن_خسته	لنگان_لنگان_رفتن_کند	لنگان_لنگان_رفتن_لنگان	راه_رفتن_با_مصدومیت	3
		s	لنگیدن_خسته	لنگان_لنگان_رفتن_آرام	لنگیدن_ارام	لنگان_لنگان_رفتن_کند	لنگیدن_آهسته	4
		rs	لنگیدن_خسته	راه_رفتن_لنگان_لنگان	راه_رفتن_لنگ_لنگان	لنگیدن_آهسته	قدم_زدن_اهسته	3
ontuu_kivuliaasti	20	-s	لنگیدن_درد	راه_رفتن_لنگان_لنگان	راه_رفتن_زخمی	لنگان_لنگان_رفتن_شل	لنگان_لنگان_رفتن_خراب	3
		s	لنگیدن_با_زحمت	لنگان_لنگان_راهرفتن_خیلی_لنگیدن	لنگیدن_خسته	لنگیدن_عاجزانه	راه_رفتن_لنگان_لنگان	4
		rs	راه_رفتن_لنگان_لنگان	راه_رفتن_لنگ_لنگان	لنگیدن_خسته	راه_رفتن_لنگان	لنگان_لنگان_راهرفتن_خیلی_لنگیدن	4
kävelee_rauhallisesti	15	-s	راه_رفتن_بی_خیال	راه_رفتن_آرام	راه_رفتن_بی_حوصله	راه_رفتن_سردرگم	راه_رفتن_گنده	1
		s	راه_رفتن_بی_حوصله	قدم_زدن_باطمانینه	راه_رفتن_آهسته	راه_رفتن_بی_هدف	راه_رفتن_آرام	3
		rs	راه_رفتن_آرام	راه_رفتن_بی_حوصله	راه_رفتن_عادی	قدم_زدن_بالنگیدن	راه_رفتن_آهسته	2
juoksee_rennosti	15	-s	دویدن_با_عجله	دویدن_تند	دویدن_با_آرامش	دویدن_بانرمی	دویدن_کمی_تند	2
		s	دویدن_معمولی	دویدن_خوشحال	دویدن_کمی_تند	دویدن_با_عجله	دویدن_با_انرژی	2
		rs	دویدن_کمی_تند	دویدن_با_عجله	دویدن_خوشحال	دویدن_تند	دویدن_معمولی	1
indexical-grounding								27
pattern-grounding								36
randomized-grounding								27

Although the number of good translations found by randomized-grounding is comparable to the number of good translations recognized by indexical-grounding, the quality of translation implemented by indexical-grounding is better because the good translations detected by this method are ranked higher. In other words, the appropriate extracted Farsi translations found by indexical-grounding are closer to the original Finnish verb-adverbs than the appropriate translations extracted by randomized-grounding. In addition, the translation of Finnish verb-adverbs to Farsi verb-adverbs is boosted by utilizing pattern-grounding.

Appendix F. Cophenetic correlation coefficient results

In the following tables, you can see the cophenetic values computed for agglomerative hierarchical clustering of English, Finnish, and Farsi verbs and verb-adverbs.

Hierarchical clustering of English verbs			
indexical-grounding		pattern-grounding	
method name	cophenet value	method name	cophenet value
average	0.8741	average	0.8928
weighted	0.8250	weighted	0.8853
complete	0.7826	complete	0.7637
single	0.7215	single	0.7593

The above table demonstrates that regardless of the selected method, pattern-grounding will lead to a clustering tree in which distances among objects reflect the original distances more appropriately than the clustering tree of the the indexical-grounding approach. In other words, when English verbs are grounded using motion data, they can be clustered more accurately.

Hierarchical clustering of English verb-adverbs			
indexical-grounding		pattern-grounding	
method name	cophenet value	method name	cophenet value
average	0.8816	average	0.8935
weighted	0.8794	weighted	0.8299
complete	0.7978	complete	0.8101
single	0.6412	single	0.7343

Cophenetic correlation coefficients show that English verb-adverbs can be clustered more precisely when pattern-grounding is applied.

Hierarchical clustering of Finnish verbs			
indexical-grounding		pattern-grounding	
method name	cophenet value	method name	cophenet value
average	0.9123	average	0.9313
weighted	0.8911	weighted	0.9181
complete	0.8095	single	0.8260
single	0.8014	complete	0.8149

The above table shows that Finnish verbs can be clustered more exactly when pattern-grounding is applied.

Hierarchical clustering of Finnish verb-adverbs			
indexical-grounding		pattern-grounding	
method name	cophenet value	method name	cophenet value
average	0.8645	average	0.9162
weighted	0.8588	weighted	0.9105
complete	0.7252	complete	0.8560
single	0.6344	single	0.5846

The above table shows that pattern-grounding can be almost always beneficial for hierarchical clustering of Finnish verb-adverbs except when single method is applied.

Hierarchical clustering of Farsi verbs			
pattern-grounding		pattern-grounding	
method name	cophenet value	method name	cophenet value
average	0.785	average	0.789
weighted	0.762	weighted	0.761
complete	0.717	single	0.719
single	0.706	complete	0.651

The above table shows that Farsi verbs can be clustered more precisely when pattern-grounding is applied.

Hierarchical clustering of Farsi verb-adverbs			
indexical-grounding		pattern-grounding	
method name	cophenet value	method name	cophenet value
average	0.889	average	0.929
weighted	0.866	weighted	0.920
single	0.797	single	0.917
complete	0.757	complete	0.840

The above table shows that Farsi verb-adverbs can be clustered more accurately using pattern-grounding.