Resolving Polysemy in Malayalam Verbs Using Context Similarity

S. N. Mohan Raj, S. Sachin Kumar, S. Rajendran, and K. P. Soman

1 Introduction

High degree of polysemy prevails in natural language, and so whatever utterance we come across is liable to be interpreted in multiple ways. But the high degree of ambiguity does not hamper our understanding of the concerned utterance. Mostly, the context invalidates the multiple interpretations and assigns a single interpretation to the given expression. It is the context which helps a native speaker to interpret an utterance or sentence correctly. Any automatic way of interpreting the sense of a lexical item expects the contexts to select or activate the correct sense out of the competitive senses of the concerned lexical items. All major word classes exhibit lexical ambiguity, and the contextual factors relevant for the concerned word resolve the meaning of the targeted word [20]. For example, adjectives are assigned meaning based on the head noun to which it is concatenated; similarly, the verb governing the noun or the modifier of the noun determines the meaning of the polysemous noun; the meaning of a verb may be determined by their argument structure or by the co-occurring elements of their syntactic frame ([24]: 215). This is illustrated below and the relevant senses are given in parentheses:

1. (a). *vEgatayuLLa kAr* "fast car" (the car that is or can be driven fast)
1. (b). *vEgatayuLLa jOlikkAr* "fast worker (one who works fast)"
2. (a). *patratte curuTTu* "role the newspaper" (physical object)
2. (b). *patratte paThikku* "read the newspaper" (content)
3. (a). *avar atinRe naSTatte uLkoNTu* "They accepted the loss" (pay)
3. (b). *avar A varttamAnam uLkoNTu* "They understood the information" (Learn)

S. N. Mohan Raj · S. Sachin Kumar (✉) · S. Rajendran · K. P. Soman
Center for Computational Engineering & Networking (CEN), Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: kp_soman@amrita.edu

R. Kumar, S. Paiva (eds.), *Applications in Ubiquitous Computing*, EAI/Springer Innovations in Communication and Computing,
https://doi.org/10.1007/978-3-030-35280-6_7

The lexical ambiguity of the verbs can be resolved by automatic means of finding sense distinctions using the semantics of the arguments of the targeted verbs. We can exploit the argument structure of a verb to interpret the correct sense of a verb. The corpus can give us the distribution of the verb with reference to its context. By means of distributional similarity, we will be able to select the correct sense of a verb. It is apparent that the resemblance of meaning between the words is revealed in resemblance of the contexts in which the words occur. It has been expressed in many ways as found in the "strong contextual hypothesis" of Miller and Charles [17] and in the well-known remark of Firth, "You shall know a word by the company it keeps" ([9]: 11).

It can be inferred that similar contexts in which a word occur reveal the similar sense of the word. This likelihood is exploited in resolving lexical ambiguity. However, making use of the notion of distributional similarity in computational applications faces problems. One of the main problems is that any kind of generalization based on distributional similarity relies on the identification of the sense in which a polysemous word occurs.

2 Types of Ambiguity

Ambiguity can emerge at any level of discourse – in words, in a sentence, or in a set of sentences. One can, thus, distinguish between referential, syntactical, and cross-textual ambiguities, depending on whether they occur in a single word, a sentence, or a set of sentences; a text. Pehar ([22]: 165–169) provided a definition and examples for each.

There are two types of ambiguity, lexical and structural. Lexical ambiguity is by far the more common. Everyday examples include nouns like *kAl* "quarter/leg," *aTa* "cake/incubation," *ara* "waist/half," and *maTi* "lap/laziness"; verbs like *Otə* "run," *aTi* "beat," and *eTu* "take"; and adjectives like *AzamuLLa* "deep," *uNakka* "dry," and *kaTTiyuLLLa* "hard." There are various tests for ambiguity. One test is having two unrelated antonyms, as with *kaTTiyuLLa* "hard," which has both *mrəduvAya* "soft" and *eLuppamuLLa* "easy" as opposites. Another is the conjunction reduction test. The test involves conjoining of two ambiguous sentences. The reduced sentence is zeugmatic. Consider the sentence, *avanu sarvakalASAlayil ninnum paTTam kiTTi; maRRonnu kaTayil ninnum kiTTi* "He received a degree from a university and another one from a shop." In the first part of the sentence *paTTam* refers to university degree and the second part it refers to "kite." Evidence that the word "*paTTam*" is ambiguous is provided by the anomaly of the "crossed interpretation" of the sentence, on which "*paTTam*" is used to refer to a university degree and "another one" to a kite.

The above examples of ambiguity are each a case of one word with more than one meaning. However, it is not always clear when we have only one word. The verb "desert" and the noun "dessert," which sound the same but, are spelled differently, count as distinct words (they are homonyms). So do the noun *muTTə* "knee" and the verb *muTTə* "knock," even though they not only sound the same but are spelled

the same. These examples may be clear cases of homonymy. But in the case of noun "*kaTi"* and the verb "*kaTi"* or the postposition *mukaLil* "over" and the adverb *mukaLil* "over," the decision of homonymy is difficult to make. We have the problem of deciding whether they are pairs of homonyms or different forms of the same word. There is no general consensus on how to draw the line between cases of one ambiguous word and cases of two homonymous words. Perhaps, the difference is ultimately arbitrary.

Sometimes, one meaning of a word is derived from another. For example, the cognitive sense of *kANuka* "see" seems derived from its visual sense. The sense of *muRi* "break," in *avan maram muRi-ccu* "He cut the tree" is derived from its sense in *maram muRi-njnju* "The tree broke." Similarly, the transitive senses of *eri "*burn," *veTTu* "*cut,"* and *poTTu* "burst" are derived from their intransitive senses. Now it could be argued that in each of these cases, the derived sense does not really qualify as a second meaning of the word but is actually the result of a lexical operation on the underived sense. This argument is plausible to the extent that the phenomenon is systematic and general, rather than peculiar to particular words. Lexical semantics has the task of identifying and characterizing such systematic phenomena. It is also concerned to explain the rich and subtle semantic behavior of common and highly flexible words like the verbs *cey* "do" and *iTu* "put" and the postpositions *uLLe* "inside," *mukaLil* "above," and *tAze* "below." Each of these words has uses which are so numerous yet so closely related that they are often described as "polysemous" rather than ambiguous.

Structural ambiguity occurs when a phrase or sentence has more than one underlying structure. For example, the phrases such as *kuLLarAya aNum peNNum* "short [men and women]," "[short man] and woman" and *veLLa marunnu kuppi* "white [medicine bottle]," "white medicine [bottle]" are structurally ambiguous. In the sentence, *kOzi tinnAn tayArAyi* "The chicken is ready to eat," that could be used to describe either a hungry chicken or a broiled chicken. It is arguable that the operative reading depends on whether or not the implicit subject of the infinitive clause "to eat" is tied anaphorically to the subject ("the chicken") of the main clause.

2.1 Lexical Ambiguity

This type of ambiguity as a rule connects to a homonym or a polysemous word presented in isolation (e.g., *aTa* "rice cake"/"blockade"/ "incubation"). Böhmerová ([4]: 28) presents the Table 1 as the typology of lexical ambiguity which is a partial representation of the numerous types of lexical ambiguity:

2.1.1 Lexical Ambiguity Due to Homonymy

If two entirely different words having same form have different meanings, the ambiguity arises due to homonymy. Homonymy is used in syntactic analysis to

Table 1 Types of lexical ambiguity

A. Language-inherent ambiguity (occurring in a particular language)
1. Onomatological (structural) ambiguity
2. Polysemy
3. Homonymy
4. Enantiosemy
B. Cross-linguistic ambiguity (occurring between or among languages)
1. Onomatological (structural) nonparallelism.
2. Lexical cross-linguistic nonparallelism (asymmetry)
3. Ambiguity due to synchronical motivational nontransparency
4. Ambiguity resulting from differing objective reality

refer to lexical items which have same form but differ in meaning. Homophones are a type of homonymy; it is used in semantic analysis to refer to words (i.e., lexemes) which have same pronunciation (e.g., *OTu* "tile"/"run"; *paTTam* "kite"/"degree"). The distinction between homographs and homophones is absent in Malayalam. When there is ambiguity between homonyms (whether nondeliberate or contrived, as in riddles and puns), a homonymic clash or conflict is said to have occurred. In semantic analysis, the theoretical distinction between homonymy and polysemy (one form with different meanings) provides a problem which has attracted a great deal of attention. Lyons ([15]: 550) identifies two kinds of lexical ambiguity, one of which depends on homonymy and the other on polysemy. He elaborately discuss about this problem giving more clarity to the concepts. He opines that the difference between homonymy and polysemy is easier to explain in general terms than it is to define in terms of objective and operationally satisfactory criteria. One of the criteria for distinguishing polysemy form homonymy is etymological information. Even etymological information may fail or unreliable sometimes. Lyons ([15]: 567) provides workable criteria based on distribution to solve the problem of homonymy and polysemy. In the following examples, the word *aTa* shows homonymy expressing contrastive sense; *aTa* can denote "rice cake" as well as "prop" or "keeping eggs for incubation."

4. (a). *avan aTa kazikkunnu* "He is eating rice cake"
4. (b). *avan vaNTiykkə aTa vaccu* "He put blockade for the cart"
4. (c). *avan muTTa aTa vaccu* "He kept the eggs for incubation"

As for machine translation is concerned both the homonymy and polysemy are treated alike as the aim is to finding out the meaning by context. The homographs belonging to different grammatical categories can be resolved as explained before. But they belong to same grammatical categories syntactic parsing may not be enough. One common approach is to assign semantic features such as *manushyan* "human," *strI* "female," and *drAvakam* "liquid" and to specify which features are compatible in the given syntactic constructions, via selection restrictions. For example, it might be specified that the verb *kuTi* "drink" have an "animate" subject.

But the expectation of animate subject for *kuTi* is challenged by the following type of usages where *kuTi* takes even inanimate subject.

5. *kARə orupATu indhanam kuTikkunnu* "The car is consuming lot of fuel"

There are difficulties in finding semantic features that can be used consistently and specifying the selection restriction for nouns and verbs based on these features. Even then, these are widely used in Machine translation system often in combination with case roles. But the semantic features cannot solve all the problems, even in situations for which they have been devised. For example, let us take the word *aTa*. As we have found out it is used in the senses of "rice cake," "keeping egg for incubation," and "prop." These three meanings can be differentiated explaining the relevant co-occurrence restrictions we find out in the following sentences in which *aTa* is used.

6. (a). *kuTTi aTa kazikkunnu* "The child is eating rice cake"
6. (b). *avan kOzhikkunjnju virikkAn vENTi irupatu muTTa aTa vaccu*

"He kept twenty eggs for incubation so as to hatch as chicken"

6. (c). *avan vaNTikku aTa vaccu* "He kept a prop for the cart"

The word *aTa* can be collocated with *kazikku* "eat" only in its eating sense; it can collocate with *vaykku* "keep" in the other two senses. *aTa* with "keeping for incubation" sense collocates with *muTTa* "egg" and with "prop" sense collocates with *vaNTi* "cart," thereby keep these two senses apart.

2.1.2 Lexical Ambiguity Due to Polysemy

If a word has two or more meanings, it can be said that the ambiguity is due to polysemy. Polysemy expresses extension of meaning. The polysemous words may express new meaning by metaphoric and metonymic extensions. For example, the word *SAkha* "branch" may denote branch of a tree as well as a branch of a bank. *naTa* can denote the action of walking as well as happening or functioning of something.

7. (a). *avan ennum rAvile sKULil naTannu pOkunnu*

"He goes to school daily by walking"

7. (b). *A stApanam nannAyiTTu naTakkunnuNTə*

"That organization is functioning well"

7. (c). *A tiyERRaRil sinima naTannukoNTirikkunnu*

"A cinema is running in the theatre"

OTu can denote human action of running as well functioning of mechanical devices.

8. (a). *avan vEgattil OTunnu* "He is running fast"
8. (b). cumar *ghaTikAram nannAyi OTunnu.*

"The wall clock is functioning well"

kaNNə may denote the eye of animate beings as well as the eye like spot in the coconut.

9. (a). *avan tanRe kaNNukaLe aTaccu* "He closed his eyes"
9. (b). *tEngaykkə mUnu kaNNukaL uNTə*

"there are three eye like spots in the coconut"

Similarly, many words denoting body parts show polysemous extension, thereby denoting ambiguity. For example, *vAy* can be a human mouth as well as the mouth of a bottle; *kAl* can be a human leg or leg of furniture. In the following example, the ambiguity is due to metonymic extension.

10. *grAmam ciriccu* "The people (of the village) laughed"

Here in this sentence, *grAmam* "people" is used as a metonymic extension of *grAmam* "village."

2.1.3 Polysemy, Homonymy, Delineating Senses

The polysemy–homonymy distinction is clear and unproblematic for the first sight. Homonyms are unrelated words that share the same spoken and written form, while a word that has two or more different, but related meanings is polysemous. The word *koTTu* is an example of polysemy, because it can refer to "a kind of percussion musical instrument," as well as "knock" and "beat." The similarity of their shape and function is reflected in their meaning; therefore, these two senses are said to be connected to the same, polysemous lexeme. Well-known examples for homonymy are *paTTam 1* "degree" and *paTTam* 2 "kite." There could be dispute over the decision of homographic words as homonyms of polysemous words. In the case of *koTTu*, the senses are related form etymological or diachronic point of view. But *paTTam 1* and *paTTam 2* cannot be diachronically or etymologically related. The *paTTam* example shows that separating polysemy from homonymy may involve diachronic considerations. However, such a strategy should be aligned with the observation that speakers of a language are more or less unaware of the etymology of words, which also means that diachronically motivated polysemy–homonymy decisions lose their psycholinguistic relevance. Etymologically *naTa* "walk" and *naTa* "happen" and *naTa* "door step" are from the same source. They are given in the dictionary as three entries. On the other hand, when the history of the language is rejected as a clue, distinguishing polysemy from homonymy may turn out to be more than challenging.

Enumeration of senses in Natural Language Processing (NLP) applications is an accepted practice, too. NLP usually resort to what Lyons calls the "maximize

homonymy" approach – by neglecting polysemy. For instance, WordNet [18], a full-scale lexical database, excludes polysemy from the description although it implements a host of other lexical and semantic relations. The presence of multiple word senses is quite typical rather than exceptional. In the Semcor corpus, for instance, Mihalcea and Moldovan [16] found 6.6 possible interpretations per word on average (using WordNet sense categorization). Even tiny sense variations are kept distinct in WordNet, and the database is probably as fine grained as possible. Mihalcea and Moldovan [16] point out that it is not uncommon that WN "word senses are so close together that a distinction is hard to be made even for humans" ([16]: 454).

2.2 Category Ambiguity

Category ambiguity is the most straightforward type of lexical ambiguity. This happens when a given word may be assigned to more than one grammatical or syntactic category as per context. One can find a number of such examples in Malayalam. For example, the word *pacca* "green" can be both noun and adverb. Similarly, the word *vEgam* can be both adverb and adverb. *kuTi* could be both verb and noun.

11. (a). *A sAriyuTe niRAm paccay-ANə* "The color of sari is green."
11. (b). *A pacca sAri vila-kkUTutal-ANə* "That green sari is costly."
12. (a). *avan vEgam vannu* "He came quickly."
12. (b). *A kArinRe vEgata kUTutal-ANə* "The speed of the car is more."

The words like *mukaLil* and *tAze* could be adverbs and postpositions.

13. (a). *avan AnayuTe mukLil kayaRi* "He climbed over the elephant."
13. (b). *avan mukaLil kayaRi* "He climbed up."
14. (a). *avan marattinRe tAze iRangi* "He climbed down from the tree."
14. (b). *avan tAze iRangi* "He climbed down."

Category ambiguities can be often be resolved by morphological inflection. For example, *kuTi* in *avan kuTikkunnu* "he drinks" is a verb and *kuTi* in *avan kuTi niRutti* "He stopped drinking" is noun. Frequently, ambiguity can be resolved by syntactic parsing. However, the problem increases when several categorical ambiguous words occur in the same sentence, each requiring being resolved syntactically.

2.2.1 Categorical Ambiguity Due to Historical Functional Reorganization

The inflected forms of nouns or verbs will denote different word category or functional category due to historical meaning change. For example, many of the postpositions in Malayalam are historically the inflected forms of verbs and nouns. The inflected forms *ninnu* "from," *paRRi* "about," *kuRiccu* "about," *koNTu* "by

(means of)," *vaccu* "by (means of)," *cuRRi* "around," *nOkki* "towards," and *kUTe* "along with" are the inflected forms of the verb *nilkkuka* "stand," *paRRuka* "apply," *kuRikkuka* "aim," *koLLuka* "have," *vaykkuka* "*to keep*," *cuRRuka* "go around," *nOkkuka* "look at," and *kUTuka* "assemble," respectively. The postpositions *munne* "before" and *pinne* "after," *munpil* "in fornt," *pinnil* "at the back," *tAze* "below," and *mukaLil* "above" are the inflected or modified forms of nouns.

15. (a). *avan vITTil ninnum vannu* "He came from home."
15. (b). *avan vITTil nilkunnu* "He is (standing) the house."
16. (a). *avan enne paRRi paRanjnu* "He talked about me."
16. (b). *avan avaLuTe manasil kayaRi paRRi* "He landed on her heart."
17. (a). *avan enne kuRiccu paRanjnju* "He talked about me."
17. (b). *avan kaTalAsil kuRiccu* "He noted in the paper."
18. (a). *avan ate katti koNTu veTTi* "He cut it by a knife."
18. (b). *avanRe kaiyil veTTu koNTu* "He wounded his hand."
20. (a). *avan katti vaccu atine veTTi* "He cut it with a knife."
20. (b). *avan katti vaccirikkunnu* "He is having a knife."
21. (a). *vITTine cuRRi marngngaLANu* "There are trees surrounding the house."
21. (b). *avan vITTine cuRRi* "He went around the house."
22. (a). *avan avaLe nOkki vannu* "He came towards her."
22. (b). *avan avaLe nOkki* "He looked at her."
23. (a). *avan avaLuTe kUTe vannu* "He came with her."
23. (b). *avarellAm aviTe kUTi* "They gathered there."

Although people are sometimes said to be ambiguous in how they use language, ambiguity is, strictly speaking, a property of linguistic expressions. A word, phrase, or sentence is ambiguous if it has more than one meaning. Obviously, this definition does not say what meanings are or what it is for an expression to have one (or more than one). For a particular language, this information is provided by a grammar, which systematically pairs forms with meanings, ambiguous forms with more than one meaning.

3 Polysemy: Its Nature and Consequences

Here, we will be discussing the nature of polysemy, consequence of polysemy, and the ambiguity due to polysemy.

3.1 Polysemy

Polysemy (/pəˈlɪsɨmi/ or /ˈpɒlɨsiːmi/ from Greek: πολυ-, poly-, "many" and σῆμα, sêma, "sign") is the capacity for a sign (such as, a word, phrase, or symbol) to have multiple meanings (i.e., multiple semes or sememes and thus multiple senses),

usually related by contiguity of meaning within a semantic field. Polysemy is thus distinct from homonymy – or homophony – which is an accidental similarity between two words (such as *bear* the animal, and the verb to *bear*); while homonymy is often a mere linguistic coincidence, polysemy is not.

Charles Fillmore and Beryl Atkins' [8] definition stipulates three elements:

(i) The various senses of a polysemous word have a central origin.
(ii) The links between these senses form a network.
(iii) Understanding the "inner" one contributes to understanding of the "outer" one.

Polysemy is a pivotal concept within disciplines such as media studies and linguistics. The analysis of polysemy, synonymy, and hyponymy and meronymy is vital to taxonomy and ontology in the information-science senses of those terms. It has applications in pedagogy and machine learning, because they rely on word-sense disambiguation and schemas.

A polyseme is a word or phrase with different, but related senses. Since the test for polysemy is the vague concept of relatedness, judgments of polysemy can be difficult to make. Because applying preexisting words to new situations is a natural process of language change, looking at words' etymology is helpful in determining polysemy but not the only solution; as words become lost in etymology, what once was a useful distinction of meaning may no longer be so. Some apparently unrelated words share a common historical origin; however, so etymology is not a dependable test for polysemy, and dictionary writers also often defer to speakers' intuitions to judge polysemy in cases where it contradicts etymology. Malayalam has many polysemous words. For example, the verb *naTa* "walk" can mean "function," "occur," "happen," etc.; *OTu* can mean "run," "function," "occur," "flow," etc. [19].

In vertical polysemy, a word refers to a member of a subcategory (e.g., *kOzi* "fowl" is for denoting "chicken"). A closely related idea is metonymy, in which a word with one original meaning is used to refer to something else connected to it. Metaphorical uses of words are instances of polysemy as well. For example, *kuTi* "drink" might be specified as "drink as animate being." This can be metaphorically extended to inanimate beings as given in the following example:

24. *I vAhanam orupATu indhanam kuTikkunnnu*

"This vehicle is consuming lot of petrol."

The difference between homonyms and polysemes is subtle. Lexicographers define polysemes within a single dictionary lemma, numbering different meanings, while homonyms are treated in separate lemmata. Semantic shift can separate a polysemous word into separate homonyms. For example, *naTa* as in *avan naTakkunnu* "he is walking" in which it originally denotes walking by means of legs. Now, it is used to denote functioning, occurring, and happening. The Malayalam lexicographer gives separate entries for these two senses. This is an example which shows how polysemy becomes homonymy. It can be inferred by psycholinguistic experiments that homonyms and polysemes are denoted differently in the mental lexicon of the people. It is often felt that the different senses of homonymous words

interfere in the comprehension, whereas the same is not true with polysemes of a word. This argument has mixed results. One group of polysemes are those in which a word meaning an activity, perhaps derived from a verb, acquires the meanings of those engaged in the activity, or perhaps the results of the activity, or the time or place in which the activity occurs or has occurred. Sometimes, only one of those meanings is intended, depending on context, and sometimes multiple meanings are intended at the same time. Other types are derivations from one of the other meanings that lead to a verb or activity.

OTu "**run**"

(a) *OTu* – move fast using legs
(b) *OTu* – work as machine
(c) *OTu* – move as time

This example shows the specific polysemy where the same word is used at different levels of taxonomy. Example ***a*** contains ***b***, and ***b*** contains ***c***.

muTTə in Malayalam shows polysemy and homonymy. *muTTə* as a verb can mean "knock," "dash," "pressure for urinating, defecating, breathing trouble, finding difficulty in answering, etc. This is an instance for polysemy. *muTTə* as a known means "knee" and "support." This is an instance for homonymy.

The verbs in Malayalam are highly polysemous as only a handful amount of verbs are used to denoted innumerable number of events, actions, and processes. The verbs such as *iTuka*, *aTikkuka,* and *OTuka* are highly polysemous showing extensions of meaning in various dimensions from the core meaning.

A lexical conception of polysemy was developed by Atkins [3] in the form of lexical implication rules. These are rules that describe how words, in one lexical context, can then be used, in a different form, in a related context. A crude example of such a rule is the pastoral idea of "verbizing one's nouns." For example, certain nouns, used in certain contexts, can be converted into a verb, conveying a related meaning (e.g., In English, the nouns *chair, bench, table,* etc. are used in verbal sense.) This kind of verbalization of nouns into verbs is not found or very rare in Malayalam. In Malayalam, certain nouns such as *praNayam* "love (noun)" become verb by some formal change, that is, *prNayikkuka* "to love (verb)." A number of nouns borrowed from Sanskrit get converted into verbs in this fashion (*lAbham* "probit" > *lAbhikkuka* "make profit").

The polysemous nature of verbs helps Malayalam to form new verbs from the combination of noun and verb. In that case, the verbs act as verbalizers.

For example, *accu* "print" + *aTi* "beat" > *accaTi* "print"
kOppi-aTi (*kOppi* "copy" + *aTi* "beat" > *kOppiyaTi* "copy"

On a scale of meaning variance, ambiguity and vagueness are the two extremes, whereas polysemy is in between the other two. It shares features with both and is a common phenomenon in everyday language use. Polysemy involves lexemes that are clearly united (share a common schema) as well as clearly separable at the same time. Polysemous words are the result of lexemes gaining new usages over time

which share the same phonological form and appear to have separate meanings to nonetymologists. *aTi* "beat" is one example of polysemy.

accaTi "print," *kaNNaTi* "wink," *kayyaTi* "clap," *muRRamaTi* "clean with broom," *ANiyaTi* "drive in a nail," *bhAgyamaTi*, "win a lottery"

Polysemy is sometimes mixed up with homonymy. However, there is a clear-cut distinction between these two. Polysemous lexemes always share the same etymological background and/or are conceived of as being semantically related by speakers, whereas homonymous words just happen to end up with the same phonological form. Therefore, homonymy may be seen as a subcategory of lexical ambiguity.

3.2 Test for Polysemy

Establishing polysemy is a challenging task. There are some tests for determining the presence of polysemy. In addition, polysemy is differentiated from other phenomena that involve potential multiplicity of meaning. A few potential cases of polysemy are to be explored. We have to deal with the polysemy paradox and consider ways in which types of polysemy can be characterized and categorized.

First, etymology is not a perfectly reliable arbiter of polysemy. A meaning that is etymologically related to another need not be appropriately similar to the initial meaning, as the two may drift apart over time to the point that they are no longer suitably related. An example of the verb *naTa* which originally denotes "moving by legs" is extended to denote "function" as in the following example:

25. *avanRe kaccavaTam nannAyi naTakkunnu*

"His business is going on well."

As seen from the above example, a word may acquire a new sense historically.

Second, it is arguable whether or not polysemy should be seen as coming in degrees. A bad argument for this conclusion goes as follows: polysemous terms enjoy related, similar meanings. Similarity comes in degrees. Therefore, polysemy comes in degrees. But the grade ability of a phrase does not follow from the grade ability of a condition for the correct application of the phrase. But whether or not polysemy comes in degrees, there are definitely cases in which it is vague whether or not two meanings for a term or phrase constitute polysemy. It will be often hard to tell whether or not a term is merely ambiguous or polysemous. The examples of *aTi*, *Otu,* and *naTa* will illustrate the above statement.

Zeugma is one of tests for polysemy. If one word seems to exhibit zeugma when applied in different contexts, it is likely that the contexts bring out different polysemes of the same word. If the two senses of the same word do not seem to fit, yet seem related, then it is likely that they are polysemous. The fact that this test again depends on speakers' judgments about relatedness, however, means that this

test for polysemy is not dependable, but is rather merely a helpful conceptual aid. The following example will illustrate this statement.

26. *avaL avanuvENTi vITinREyum hrudayattinREyum vAtil tuRanniTTu*

"She waited leaving open her house door and heart"

Linguists and philosophers have developed various tests for polysemy and ambiguity. Most of these tests involve attempts to "freeze" a single meaning of the putatively polysemous phrase and then see if that frozen meaning can be used to express the multiple meanings. Unfortunately, this makes most tests unsuitable for distinguishing polysemous phrases from merely ambiguous phrases, since both enjoy multiple meanings.

These are tests, not knockdown arguments. Most of them require a somewhat slippery notion of interpretation. One may wonder whether the sentence in question expresses multiple meanings or not. Of course, with some creativity and perhaps some practice, competent language users can find themselves detecting meanings in unexpected places. What the tests presuppose is that the reader can access something like interpretations of a sentence that are not deviant.

4 Resolving Polysemy

This is the main focus of the chapter. Here, we will be discussing about the resolving polysemy by contextual similarity.

4.1 Relevance of Context for Resolving Polysemy

According to Rumshisky ([24]: 217), the combination of the following two factors assigns meaning to a word: the syntactic frame of a word and the semantics of the word in that syntactic-dependent frame. Such words are referred as "selector" by Rumshisky ([24]: 217). This is applicable for both the head word and dependent word. The syntactic frame is extendable to the minor categories such as adverbials, locatives, and temporal adjuncts and to the subphrasal cues such as genitives, partitives, negatives, bare plural/determiner distinction, and infinitives ([24]: 217). The set of all "usage contexts" in which a polysemous word occurs can usually be divided into groups. Each group roughly corresponds to a distinct "sense."

Consider following sentences with the verbs *niSEdhikkuka* "deny" made use of in 27*a* and *paRayuka* "say" made use of in 28*a* to illustrate the contribution of different context parameters to disambiguation. The difference in the syntactic patterns for the verb *niSEdhiccu* "deny" as shown in 27*b* and 27*c* disambiguate between the two dominant senses: "refuse to grant" and "proclaim false" and similarly the difference in the syntactic patterns of *paRayuka* "say" as shown in 28*a* and 28*b* disambiguate between "complain" and "blame."

Syntactic frame

27. (a). *adhikAri atinu kAraNam uNTenna kAriyatte nishEdhiccu* {*enna*-clause)

"The authority denied that there is a reason for that" (proclaim false)

27. (b). *adhikAri kAraNatte nishEdhiccu* [NP]

"The authority denied the reason." (proclaim false)

27. (c). *adhikAri rAdhaykku visa nishEdhiccu* [NP]

"The authority dined visa to Radha." (refuse to grant)

28. (a). *avaL koccine aTiccu ennU avan kuRRam paRanjnu* [*ennu*-clause]

"He complained that she bet the child." (complain)

28. (b). *avan avaLe kuRRam paRanjnju* (blame) [NP]

"He blamed her."

Consider the following sentences with *OTuka* "run," *valicceTukkuka* "absorb," and *perumARuka* "behave." The contrasting argument and/or adjunct semantics shown in 2*9, 30,* and *31* evoke the different the senses of *OTuka* "run," *valicce-Tukkuka* "absorb," and *perumARuka* "behave," respectively. The relevant argument type is shown in brackets and the corresponding sense is shown in parentheses. The following examples show the semantics of the arguments and adjuncts/adverbials:

29. (a). *pOlIs kaLLanRe puRakil OTi*

"The police ran after the thief." [chased person] (run)

29. (b). *avaL avanRe kUTe OTippOyi*

"She ran away with him." [accompanying person] (elope)

30. (a). *panjnji veLLatte valicceTuttu*

"The cotton absorbed the water." [water] (absorb}

30. (b). *avaL avnRe kaiyil ninnum paisa valicceTuttu* [money] (take away)

"She extracted money from him."

31. (a). *avan avaLOTU mariyadayAyi perumARi*

"He behaved with her with respect." [with respect] (behave)

31. (b). *pOlIs avane nallavaNNam perumARi*

"Police beat him severely." [severely] (beat)

A lexicographer establishes a set of senses available to a particular lexical item and (to some extent) specifies the context elements which typically activate each sense. This procedure is formalized in several current resource-oriented projects

such as FrameNet and Corpus Pattern Analysis (CPA). The semantics of the arguments is represented in FrameNet by making use of Fillmore's case roles. Hanks and Pustejovsky [12] attempt to arrange prototypical norms of usage for individual words in terms of contextual patterns by CPA.

4.2 Polysemy and Distributional Similarity

A number of tasks in NLP make use of the notion of distributional similarity. Distributional similarity is exploited in the areas such as word-sense disambiguation (WSD), sense induction, automatic thesaurus construction, selectional preference acquisition, and semantic role labeling ([24]: 220). Semantically similar words (as in thesaurus construction) or similar uses of the same word (as in WSD) and sense induction are identified by making use of distributional similarity. Distributional similarity results in clusters of distributionally similar words. The problem of data sparsity faced by many NLP tasks is addressed by these clusters. The distributional similarity of a polysemous word gives rise to the generalization that must be applied to different "senses" rather than to its entire occurrence regularly. The external knowledge sources such as FrameNet, machine-readable dictionaries, and WordNet give rise to information that represents semantics of the arguments ([24]: 220).

A feature vector represents a context. Each feature vector is mapped to some context component. The target word along with the context component gives the frequency of occurrence of the value of each feature vector. All the context structures or the probability of co-occurrence distribution represent the feature vector. Different approaches make use of different methods. Widdows and Dorrow [28] and Agirre et al. [1] make use of co-occurrence graph. Schütze [26], Gale et al. [10], and Widdows and Dorrow [28] make use of "distributional features based on bag-of words style co-occurrence statistics." Grefenette [11], Lin [14], and Pantel and Lin [21] make use of context structures incorporating syntactic information and sometimes semantic information from external sources.

From the distributional exemplification of the target word, the occurrences corresponding to each sense can be separated out. This will lead to the resolving of polysemy. According to Schütze [26], Grefenette [11], and Lin [14], clustering similar occurrence context for each word or clustering of actual words which are distributionally similar resolves this problem.

4.3 Sense Assignment for Polysemous Verbs

Sense inventories for polysemous verbs are often comprised by a number of related senses. Computational approaches to word-sense disambiguation assign a sense to each word in an utterance from an inventory of senses. This simplified statement may not be true when the meaning of a complex utterance is computed. Take,

for example, a target verb showing polysemy with certain semantic inclinations. Different sematic components select different senses of a verb in a given argument position. In the sentence 2*9b*, the "elope" sense of *OTuka* is selected by the argument indicating the accompanying person while "run," sense in 6a is selected by the argument indicating the "chased person"; similarly, in *30b*, the "take away" sense is selected by the argument denoting "money" in the direct object position, whereas the "absorb" sense in *30a* is selected by the argument denoting the direct object "water." In the same way in *31b*, the "beat" sense is selected by the adverbial adjunct "*nallavaNNam*" "severely" and the "behave" sense in *31a* is selected by the adverbial adjunct *maryAdayAyi* "with respect."

4.3.1 Sense-Activating Argument Sets for a Polysemous Verb

The same meaning of a verb can be triggered by a number of semantically different arguments. Certain pertinent semantic feature will be central to the interpretation of the meaning for some of the verb. Other verbs are allowed merely by a contextual interpretation. An ad hoc semantic category will be prompted by each sense of the target verb in the relevant argument position. For example, consider the senses of the verb *eTukkuka* "take (something)," "raise (hood like a snake), "take photo," "score," "copy," etc. The following examples exemplify the lexical items that occur in the position of direct object.

32. (a). *pAmpU paTam eTuttu* "The snake raised its hood."
32. (b). *avan avaLuTe paTam eTuttu* "He photographed her."
32. (c). *avan atinRe pakarppU eTuttu* "He took the copy of it."
32. (d). *avan kaNakkinU nURu SatamAnam mArkkU eTuttu*

"He scored hundred percent mark in mathematics."

The verbs need to be activated or evoked for a relevant sense by its arguments. The semantically quite discrete nouns in each argument set activate the relevant sense of the verb. A particular aspect of the sense is selected by the context provided by the verb. Mostly, an argument set carry a central component of their meaning as well as certain other peripheral components. The argument set has core members which are polysemous. In order to activate the appropriate sense of the verb, a "bidirectional selection" process needs to be implemented. But notice that the interpretation of *33a* and *33b*, for example, is quite different.

33. (a). *pOlIs avane kasRRaTiyil eTuttu* "The police arrested him."
33. (b). *avan A jOli ceyyAn kuRaccu samayam eTuttu*

"He took some time to do the work."

In the above sentences, both the words in the argument position activate the same sense of *eTukkuka*. But the disambiguation is between the EVENT reading and the TIME reading. It has to be noted that the diverse dependencies the word enters into

rely on the different aspects of the meaning. For example, consider the use of the noun arguments with the verb *eTukkuka* in the sentences *34a* and *34b* given below.

34. (a). *kamsanRe kArAgrhattil krishNan janmam eTuttu*

"Krishnan was born in Kamsan's jail."

34. (b). *innale kOTatiyil ninnum avane jAmyattil eTuttu*

"He has been bailed out from the court."

In the above examples, the words *janmam* "birth" and *jAmyattil* "bail-in" activate different senses for *eTukkuka* "take." The context provided by the verb effectively changes the relevant semantic components in the interpretation of the senses.

4.3.2 Sense Separation Based on Selector

In the case of homonymy, sufficiently distinctive semantic features are chosen for denoting different senses ([24]: 224). The relevant lexical items can be grouped together by making use of the overall distributional similarity between arguments. For example, movement sense of *maTangnguka* "return" is easily differentiated from the cluster of senses associated to "fold as paper."

35. (a). *avaL OphIsil ninnum vITTilEkkU maTangi*

"She returned from office."

35. (b). *pustakattinRe pEjukaLokke maTngippOyi*

"The pages of the book got folded up."

Similarly, the movement sense of *naTakkuka* "walk" is easily differentiated from the cluster of senses associated to "go on."

36. (a). *kunjnju patukkeppatukke naTakkAn tuTangi*

"The child started walking slowly."

36. (b). *aviTatte tiyETTaRil nalla cinima naTakkunnu*

"A cinema is going on in the theatre."

In the case of polysemy, differentiating dissimilar meanings of the verb is extremely difficult. This problem has been the subject of extensive study in lexical semantics. It aims at addressing the question of selecting distinct senses based on context. There is no clear-cut method to say when a context selects a distinct sense or when a context simply modulates the meaning [2, 7, 23]. This is crucial for the computational method of word-sense disambiguation. Lexicographers often face problems in determining when to define a set of usages as a distinct sense while "lumping and splitting" senses during dictionary construction. The sense separation is frequently determined on ad hoc basis. It results in instances where the same

occurrence may fall under more than one sense category simultaneously resulting in numerous cases of "overlapping senses." Resolving verbal polysemy often runs into this problem of indecision.

There is a necessity to decide what sense is activated by which selector. While resorting to such attempt, we should have in mind that at least some of the meanings of the verb are interconnected. The incidences of a polysemous verb in a corpus cluster into certain number of groups, each roughly corresponding to a sense ([24]: 225). There are a lot of instances in which the meaning dissimilarities are straightforward and easily noticeable. But there are some boundary cases where the sense diction of the verb is not clear cut. According to Rumshisky ([24]: 225), three kinds of selectors are possible in a given argument position: (i) Good disambiguators: Here, one meaning of the target word is instantaneously selected by the selector. (ii) Poor disambiguators: Here, the selectors may choose either sense; more contexts may be needed to disambiguate themselves. (iii) Boundary cases: Here, it is impossible to make the choice between two senses of the target word. For instance, for the verb *kANikkuka*, "show" in *37*, *sarve* "survey," and *paTam* "photo" are good disambiguators, whereas g*rAph* "graph" is a clear example of a boundary case.

37. (a). *A paTam avaLuTe mughatte nalla vaNNam kANikkunnu* ("picto rially represent")

"That picture shows her face well."

37. (b). *A sarve vyvasAya mEkhalayil uLLa sarkkArinRe purOgatiye kANikkunnu*

"The survey shows the improvement of government in industrial sector." ("demonstrated by evidence or argument")

37. (c). *I grAph I mAsattil uLLa maZyuTe SarASari aLavine kANikiunnu*

"The graph shows an average rainfall in this month." (both senses?)

Each individual sense needs to be clearly defined for the identification of boundary cases. Such instances are better interpreted as examples of "multiple selection" (i.e., simultaneous initiation of both meanings); they are not simply the examples of definitions of overlapping meaning [25]. Even syntactic structure cannot always supersede the elucidation inherent to some selectors. For instance, in *38*, it is almost impossible to resolve *nishEdhikkuka* "deny" between "refuse to grant" and "proclaim false."

38. (a). *vayOdhikarkkU avaruTe stAnam nishEdhikkappeTunnu*

"Elders are often denied the status of adulthood."

38. (b). *cila jAdikkAr strIkaLkkU svAtandryam nishEdikkunnu*

"People of certain caste denies autonomy to women."

Instead, in *39*, the selector itself is polysemous, with two elucidations obtainable for it, and there is a requirement for disambiguation by context before it can trigger the suitable meaning of the verb.

39. (a). *paNTatte aphiprAytte nishEdikku*

"Deny the traditional view." ("proclaim false")

39. (b). *avanRe anuvAdam nishEdikku*

"Deny him permission." ("refuse to grant")

The succeeding sections elaborate on the scheming of a computational approach for automatic meaning detection.

4.4 Contextual Similarity and Dissimilarity

By means of contextual similarity and dissimilarity, we can assign senses for a verb automatically text. Based on the specific task, the decision that one context is similar to the other may vary. If our aim is to find out the meaning of a verb based on its arguments, it can be inferred that different contextual arguments give different senses for the concerned verb.

For instance, the verb *eTukkuka* "take" and the direct object relation define a particular context of incidence for the noun. At its most basic, contextual similarity between occurrences of the concerned verb should reflect to what extent the contexts in which the verb occur will overlap. Similarity between contexts may be conveyed by the similarity of their arguments.

40. (a). *avan kuTTiye kaiyil eTuttu* "He took the child in his hand."
40. (b). *avan kalline kaiyl eTuttu* "He took the stone in hand."
41. *avan kaNakkinu nURu mArkku eTuttu*

"He scored hundred marks."

It can be inferred from the sentences *40a* and *40b* that the word *eTu* occurring in the two instances overlap in the context of their occurrence, thereby inferring the same sense. In sentence *41, eTu* occurs in a dissimilar context, thereby inferring a different sense.

4.5 Algorithm Architecture

The similarity and dissimilarity between contexts is made use of here to separate the senses of a verb. But, computing similarity and dissimilarity between contexts poses a problem. Two or more contexts must be similar with respect to their selectional

properties, that is, the same semantic components in the specified argument position. Similarity and dissimilarity between contexts are achieved in the present work of WSD by neural network approach.

4.6 Neural Network Approach

Automated language understanding requires the determination of the concept which a given use of a word represents. This process is referred to as word-sense disambiguation (WSD). WSD is typically affected in natural language processing systems by utilizing semantic feature lists for each word in the system's lexicon, together with restriction mechanisms such as case role selection. However, it is often impractical to manually encode such information, especially for generalized text where the variety and meaning of words is potentially unrestricted. Furthermore, restriction mechanisms usually operate within a single sentence, and thus, the broader context cannot assist in the disambiguation process.

WSD is executed using neural network approaches suggested by Cottrell and Small [6] and Waltz and Pollack [27]. The nodes ("neurons") representing words or concepts connected by "activatory" links make networks which are found in these models: the words activate the concepts to which they are semantically related and vice versa. The "lateral" inhibitory links usually interconnect competing senses of a given word. The nodes corresponding to the words in the sentence to be analyzed are activated initially. The adjacent ones of these words are activated in the next cycle in turn; then, the immediate adjacent ones of these neighbors are stimulated, and so on. Using a parallel, analog, relaxation process, the network stabilizes in a state in which one sense for each input word is more activated than the others after a number of cycles.

Neural network approaches to WSD seem able to capture most of what cannot be handled by overlap strategies such as Lesk's [13]. However, the networks used in experiments so far are hand coded and thus necessarily very small (at most, a few dozen words and concepts). Due to a lack of real-size data, it is not clear that the same neural net models will scale up for realistic application. Further, some approaches rely on "context setting" nodes to prime particular word senses in order to force the correct interpretation. But as Waltz and Pollack point out, it is possible that such words are not explicitly present in the text under analysis, but may be inferred by the reader from the presence of other related words.

To solve this problem, words in such networks have been represented by sets of semantic "micro-features" [5, 27] which corresponds to fundamental semantic distinctions, characteristic duration of events, locations, and other similar distinctions that humans typically make about situations in the world. Each concept in the network is linked, via bidirectional activatory or inhibitory links, to only a subset of the complete micro-feature set. A given concept theoretically shares several micro-features with concepts to which it is closely related and will, therefore, activate the nodes corresponding to closely related concepts when it is activated itself.

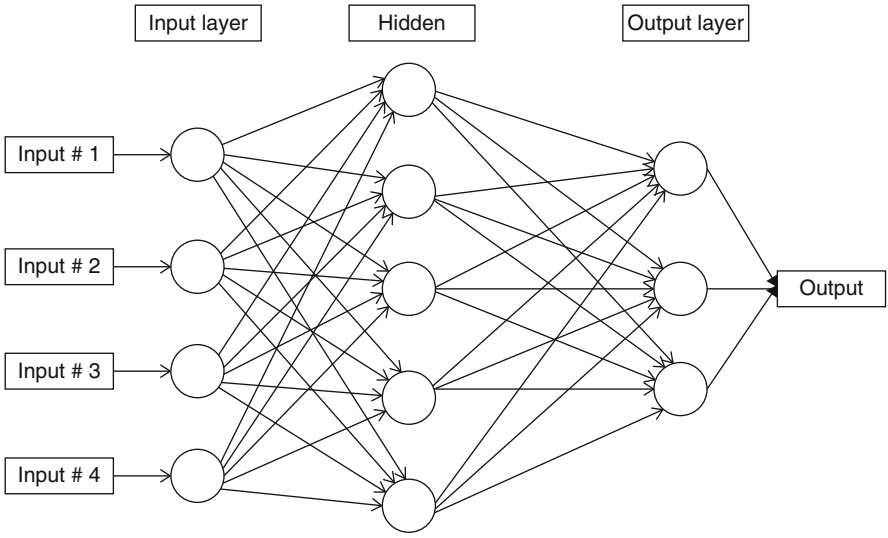

Fig. 1 Neural network disambiguation example

The words are complex units in the model we have adopted. Each word in the input is represented by a word node connected by excitatory links to sense nodes representing the different possibilities. Each sense node is in turn connected by excitatory links to word nodes representing the words in tile definition of that sense. This process is repeated a number of times, creating an increasingly complex and interconnected network. In the network, all words are reduced to their lemmas and grammatical words are excluded. The different sense nodes for a given word are interconnected by lateral inhibitory links (Fig. 1).

When the network is run, the input word nodes are activated first. Then each input word node sends activation to its sense nodes, which in turn send activation to the word nodes to which they are connected, and so on throughout the network for a number of cycles. At each cycle, word and sense nodes receive feedback from connected nodes. Competing sense nodes send inhibition to one another. Feedback and inhibition cooperate in a "winner-take-all" strategy to activate increasingly related word and sense nodes and deactivate the unrelated or weakly related nodes. Eventually, after a few dozen cycles, the network stabilizes in a configuration where only the sense nodes with the strongest relations to other nodes in the network are activated. Because of the "winner-take-all" strategy, at most one sense node per word will ultimately be activated.

4.7 Analysis of Annotation Decisions

In order to identify the senses of a verb, say, for example, *OTu* "run," a simple recurrent neural network-based learning approach is applied. It is from the human

understanding of word from its previous word in the sequence recurrent neural network (RNN) takes its idea. For the understanding of more about the present events based on the previous events, earlier neural network did not have the mechanism. The loops of neural networks which are RNN allow the information to exist in it. A simple RNN model is shown in Fig. 2.

In Fig. 2, "NN" denotes any neural network architecture, "$\mathbf{x_t}$" denotes an input, and "$\mathbf{h_t}$" denotes the output. Output value to the next phase is passed by loop. The present information with the immediate previous value is captured essentially with the help of RNNs. However, if the immediate previous value or values do not contribute for the understanding of the sequence of words, it will be omitted. An extended version of RNN long short-term memory is used in such cases, which can even capture long dependencies. In order to distinguish the nine sense classes of *OTu*, a simple RNN is used.

The evaluation measures acquired for nine epochs are shown in Tables 2 and 3. The network can be modified to get better results by creating more words and its corresponding senses.

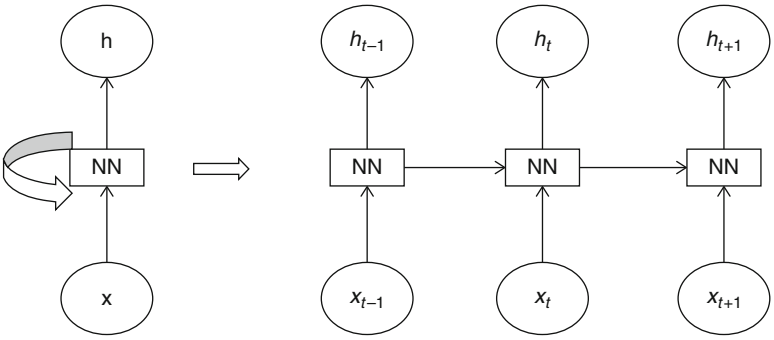

Fig. 2 Simple RNN

Table 2 No. of data in each class

Sense No.	No. of data
1	1001
2	1099
3	1045
4	1057
5	1100
6	1568
7	1154
8	1326
9	1049

Table 3 Result of nine epochs, each epoch is run for 1000 steps

Accuracy	Precision	Recall	F-measure
0.34328	0.368	0.343	0.332
0.69403	0.765	0.694	0.701
0.93284	0.938	0.933	0.933
0.97015	0.972	0.97	0.97
0.97761	0.981	0.978	0.978
0.85075	0.876	0.851	0.856
0.90299	0.905	0.903	0.902
0.98507	0.985	0.985	0.985
0.99254	0.993	0.993	0.993

5 Conclusion

The result of the proposed method is encouraging as shown Tables 1 and 2. Each of the target senses is associated with a cluster of selectional equivalents for that sense. The selectional equivalents are presented as contextualized vectors of dependable selectors. The clusters resulted from this are used to find out the selectors that trigger each sense. The association score obtained for each selector indicates which sense it tends to trigger. We are able to resolve some of the problems faced by the previous efforts to resolve polysemy. The evaluations discussed in the result section show that simple RNN provides a competing result for identifying the senses of *OTu* "run." This approach can be improved if we have a huge corpus covering all the senses of verbs. Such corpus is very difficult to get. Malayalam does not have a corpus which can be made use of for NLP works such as WSD. If such corpus is available, we can automatically classify the different senses of a verb by means of contextual clusters. We are able to identify the possible number of senses for a verb automatically.

References

1. Agirre, E., et al. (2006). *Two graph-based algorithms for state of the art WSD*. Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing 22–23 (July 2006), Sidney, Australia, pp. 585–593.
2. Apresjan, J. U. (1973). Regular polysemy. *Linguistics, 142*(5), 5–32.
3. Atkins, S. (1991). Tools for computer-aided corpus lexicography: The Hector project. *Acta Linguistica Hungarica, 41*, 5–72.
4. Böhmerová, A. D. A. (2010). Lexical ambiguity as a linguistic and lexicographical phenomenon in English, with comparisons to Slovak. In J. Kehoe (Ed.), *Ambiguity, Conference Proceedings* (pp. 27–35). Verbum, Ružomberok.
5. Bookman, L. A. (1987). *A microfeature based scheme for modeling semantics*. Proceeding of the 10th International Joint Conference on Artificial Intelligence, IJCAI'87, Milan, Italy, pp. 611–614.
6. Cottrell, G. W., & Small, S. L. (1983). A connectionist scheme for modelling word sense disambiguation. *Cognition and Brain Theory, 6*, 89–120.

7. Cruse, D. A. (1995). Polysemy and related phenomena from a cognitive linguistic viewpoint. In *Computational lexical semantics*. Cambridge: Cambridge University Press.
8. Fillmore, C. J., & Atkins, B. T. S. (1991). *Invited lecture presented at the 29th Annual Meeting of the Association for Computational Linguistics*, 18–21 June 1991, Berkeley, CA.
9. Firth, J. R. (1957). *Papers in linguistics 1934–1951*. Oxford: Oxford University Press.
10. Gale, W. A., Church, K. W., & Yarowsky, D. (1993). A method for disambiguating word senses in a large corpus. *Computers and the Humanities, 26*, 415–439.
11. Grefenette, G. (1994). *Exploration in automatic thesaurus discovery*. Norwell, MA: Kluwer Academic Publishers.
12. Hanks, P., & Pustejovsky, J. (2005). A pattern dictionary for natural language processing. *Revue Française de Linguistique Appliquee, 10*(2).
13. Lesk, M. (1986). Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. In V. DeBuys (Ed.), *Proceedings of the 5th Annual international Conference on Systems Documentation (Toronto, ON, Canada)*. SIGDOC, June 1986, pp. 24–26.
14. Lin, D. (1998). *Automatic retrieval and clustering of similar words*. Proceedings of the 17th International Conference on Computational Linguistics, Montreal, Canada.
15. Lyons, J. (1977). *Semantics volumes I and II*. Cambridge: Cambridge University Press.
16. Mihalcea, R., & Moldovan, D. (2001). *EZ.WordNet: Principles for automatic generation of a coarse grained WordNet*. Proceedings of FLAIRS, Key West, FL, pp. 454–459.
17. Miller, G. A., & Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language, Cognition and Neuroscience, 6*(1), 1–28.
18. Miller, G. A., Beckwith, R., Fellbaum, C. H., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography, 3*(4), 235–244.
19. Mohan Raj, S. N., Sachin Kumar, S., & Rajendran, S. (2017). Resolving polysemy in Malayalam verbs. *Language in India* wwwlanguageinindia.com ISSN 1930-2940, 17:1, pp. 153–163.
20. Mohan Raj, S. N., Sachin Kumar, S., Rajendran, S., & Soman, S. K. P.. (2019). Word sense disambiguation of Malayalam nouns. *Recent advances in computational intelligence*. ISSN 1860-949X ISSN 1860-9503 (electronic) *Studies in computational intelligence*. ISBN 978-3-030-12499-1 ISBN 978-3-030-12500-4 (eBook). Springer Nature Switzerland AG 2019.
21. Pantel, P., & Lin, D. (2002). *Discovering word senses from text*. Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 613–619.
22. Pehar, D. (2001). Use of ambiguities in peace agreements. In J. Kurbalija & H. Slavik (Eds.), *Language and diplomacy* (pp. 163–200). Malta: DiploFoundation.
23. Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
24. Rumshisky, A. (2008). *Verbal polysemy resolution through contextualized clustering of arguments* (Brandeis University Ph.D).
25. Rumshisky, A., Grinberg, V. A., & Pustejovsky, J. (2007). Detecting selectional behavior of complex types in text. In B. Pierrette, L. Danlos, & K. Kanzaki (Eds.), *Fourth international workshop on generative approaches to the lexicon*. Paris, France.
26. Schütze, H. (1998). Automatic word sense discrimination. *Computational Linguistics, 24*(1), 97–123.
27. Waltz, D. L., & Pollack, J. B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science, 9*, 51–74.
28. Widdows, D., & Dorrow, B. (2002). *A graph model for unsupervised lexical acquisition*. 19th International Conference on Computational Linguistics, August, 2002, Taipei, Taiwan, pp. 1093–1099.