CHAPTER 2: Review of Literature

In this Chapter, some of the main approaches dealing with the phenomenon ‘polysemy’ are reviewed. These approaches are Traditional Approach (Lyons 1977; Palmer 1981; Cruse 1986), Cognitive Linguistics (Johnson 1987; Lakoff 1987; Taylor 1995), Lexical Semantics and the Generative Lexicon (Pustejovsky 1995), Nunberg’s Pragmatic Theory of polysemy (Nunberg 1979), Kilgarriff’s approach (1992), systematic versus non-systematic polysemy (Apresjan 1973), Deane’s theory of polysemy (Deane 1987), Principled polysemy approach (Evans 2005), Natural Semantic Meta-language ‘NSM’ (Wierzbicka 1996), Relational approach (Fellbaum 1990) and the Lexicographers’ treatment of polysemy.

2.1. Historical background of polysemy

The complex relations between meanings and words were first noted by the Stoics. However, concrete research into the multiplicity of meaning only began in the 18th century and was continued in the nineteenth century by ‘linguists interested in meaning from the point of view of etymology, historical lexicography or historical semantics’ (Nerlich & Clarke, 1997: 351).

An important linguist in this nineteenth century historical tradition was Breal, whose research into polysemy marked a new starting point, in that he shifted the study of polysemy away from lexicography and etymology and
investigated ‘polysemy as the always synchronic pattern of meanings surrounding a word, which is itself the ever changing result of semantic change’ (Nerlich & Clarke, 1997: 378). The focus of studies on polysemy shifted from diachronic perspective to synchronic perspective.

With the emergence of cognitive linguistics in the eighties, the concern for relationship between language and psychology has grown and the notion that lexical items are conceptual categories, that have to be studied and investigated as reflecting general cognitive principles rather than purely formal linguistic principle, has penetrated through to linguists. Their interests toward polysemy increased and polysemy became an essential issue in linguistics.

2.2. Traditional approach

The traditional distinction between polysemy and homonymy is based on whether there is one or two lexical items involved. Lyons (1977: 550) considers them as two types of lexical ambiguity and introduces some criteria for deciding when it is polysemy and when it is homonymy.

One criterion is etymological information about the lexical item in question. Lexical items with the same origin are considered as polysemic, whereas if they have evolved from distinct lexemes in some earlier stage of the language then they are regarded as homonymous. This condition is neither satisfactory nor decisive because the history of the language does not always reflect its present state. For instance, in present-day English, the lexemes pupil,
‘student’ and *pupil* \(_2\) ‘iris of the eye’ are not usually related by native speakers, but they are both derived from Latin pupillus/pupilla ‘ward, orphan-boy’ which is itself a diminutive of *pupus* ‘child’. The opposite case is also fairly common, namely when native speakers consider two lexemes derived from different roots in an earlier stage of the language as related. For example, the lexemes *ear* \(_1\) ‘organ of hearing’ and *ear* \(_2\) ‘spike of corn’ come from two different origins: *ear* \(_1\) evolves from OE ëare from IE *aus-* (cf. Latin auris ‘ear’) and *ear* \(_2\) from OE ëar (cf. Latin acus, aceris ‘husk’) and they merged into er(e) in ME. However, most people nowadays treat these two lexemes as one polysemous word and explain their relation by means of metaphor. Therefore, the etymological criterion can be very misleading when deciding between homonymy and polysemy.

Another criterion is the unrelatedness vs. relatedness of meaning; i.e. the native speaker’s feeling that certain meanings are connected and that others are not. One of the major drawbacks that Lyons states for this criterion is that relatedness of meaning appears to be a matter of degree, together with the fact that sometimes native speaker’s intuitions are far from being the true interpretation, as has been seen with the ear example above. Attempts to formalize this relatedness of meaning have also been made. Katz (1972), Katz and Fodor’s (1963) Componential Analysis proposes the decomposition or breakdown of the sense of a word into its minimal distinctive features, i.e. into semantic components which contrast with other components. These minimal distinctive
features produce formulae called componential definitions of the type [± human], [± adult], [± male] for the description of lexemes such as man, woman, girl, boy in the semantic field of ‘human race’ (see Leech 1981: 96ff.).

Unfortunately, this type of approach is not sufficient for the polysemy-homonymy problem. First, the relatedness in the different sense of a word might not be expressible in terms of ± features and also because in some cases, these features are present in different degrees, not in absolute terms. A classical example of this problem is the word bachelor (Fillmore 1997, 1982). In a simplified world, where people are marriageable at a certain age, mostly marry at that age and stay married to the same person, bachelor is just any unmarried male past marriageable age. However, outside this simplified world, the word bachelor does not apply. That is why we find it so odd to call the Pope or a twice-married divorce bachelor, even though they both meet the criteria of the definition given above. Secondly, as Lyons (1977: 553) points out, “the possibility or impossibility of decomposing the senses of lexemes into a (structured or unstructured) set of semantic components is irrelevant, unless we can specify just how many components, or alternatively what kind of components, two senses must share in order for them to meet the criterion of relatedness of meaning”.

A third way of attempting to establish polysemy is to search for a central or core meaning. Based on the classical definition of a category as a set of necessary and sufficient conditions for membership, Allerton (1979) proposes that
when different senses of a lexeme share a core meaning, they are polysemous. On the other hand, cases when the core meaning cannot be extracted are to be considered as homonymous. For instance, the word *paper* can mean ‘newspaper’, ‘document’ and ‘academic lecture’; all these senses share the core meaning of ‘important written or printed material’. According to Palmer (1981: 105), this is possible when we have cases of metaphors and the other senses have been transferred from that core meaning. The disadvantage of this criterion is again to decide what the core meaning is. Under the Cognitive Linguistics approach, neither the core meaning approach nor Palmer’s acceptance of it in metaphorical cases is accepted. The reason is the fact that metaphor is understood as a motivated transfer between two different domains and this core meaning approach totally defeats any attempt to show a motivated account of semantic extension. The alternative to this approach within Cognitive Linguistics is the ‘family resemblance model’ (Taylor 1995: 106) or what Lakoff (1987: Ch. 6, 1996: Ch. 1) calls ‘radial categories.

Hence, the traditional approach defines polysemy as the case when a lexical item has a range of different meanings. Polysemy can be differentiated from homonymy by using a set of criteria, such as the etymology, the unrelatedness of meaning, the central or core meaning as well as some ambiguity tests. It has been argued that this model is mainly concerned with a descriptive
analysis of polysemy, without addressing questions such as why and how polysemy is created.

2.3. Cognitive approach

In the realm of cognitive linguistics, studies of polysemy are concerned with categorization. Lakoff (1987: 5) emphasizes the importance of categorization as can be seen in the following quotation:

*Categorization is not a matter to be taken lightly. There is nothing more basic than categorization to our thought, perception, action, and speech. Every time we see something as a kind of thing, for example, a tree, we are categorizing.*

Conception of categorization can be traced back to Aristotle. Lakoff, however, asserts that the traditional categorization is not suitable for illustrating polysemy.

The classical theory of categories does not do very well on the treatment of polysemy. In order to have a single lexical item, the classical theory must treat all of the related senses as having some abstract meaning in common—usually so abstract that it cannot distinguish among the cases and so devoid of real meaning that it is not recognizable as what people think of as a meaning of a word. And where there are a large number of related senses that don’t all share a property, then the classical theory is forced to treat such cases as homonymy (Lakoff, 1987: 416).
Within this framework, the main distinction between polysemy and homonymy is the systematic relation of meanings that takes place in polysemy (Lakoff 1987: 316; Johnson 1987: 193). When speaking about polysemy, the fact that we are dealing with multiple meanings is not the main point but the fact that those multiple meanings are related in a systematic and natural way.

According to Lakoff (1987), polysemy has to be understood as categorization, that is to say the idea that related meanings of words form categories and that these meanings bear family resemblance, an idea introduced by Austin (1961). Taylor (1995: 108) explains this family resemblance category in terms of ‘meaning chains’. A lexeme can convey different meanings, A, B, C, D, …A is related to B in virtue of some shared attribute(s) or other kind of similarity. Meaning B in turn becomes the source of a further extension to meaning C and so on. This ‘meaning chain’ can be represented in the following way where any node in a meaning chain can be the source of any number of meaning expressions:

A → B → C → D...

Taylor compares these ‘meaning chains’ to Lakoff’s ‘radial categories’. A category is structured radially with respect to a number of subcategories: there is a central subcategory, defined by a cluster of covering cognitive models and in addition, there are noncentral extensions which are not specialized instances of the central subcategory, but variants of it. The extensions of the central model are
not random, but motivated by the central model plus certain general principles of extension. One of the advantages of this approach if compared with classical models is that it offers adequate means of characterizing the situations where one or more senses are central or more representative.

Polysemy is the result of the extension of Idealized Cognitive Models (ICMs) to form radial categories. Sometimes, a single ICM can be the basis for a collection of senses that form a single natural category. For instance, the ICM of the lexeme *window* can take three meanings: ‘an opening in the wall’, ‘a frame fitting into the wall’ and ‘the glass filling the frame fitting into the wall’. These three senses are not unrelated; they form a natural category of senses, where correspondences remain physical. These correspondences have been explained in terms of ‘image schemata’; i.e. recurring structures of, or in, our perceptual interactions, bodily experiences and cognitive operations (Johnson 1987: 79).

In some other cases, these correspondences do not take place within the same ICM, but between the ICMs of two domains. Lakoff and Johnson (1980) propose ‘conceptual metaphor’ as one of the means for relating the different senses of a word. Metaphor is understood as an experientially-based mapping from an ICM in one domain to an ICM in another domain.

For an alternative theory of categorization, Lakoff suggests *prototype theory*. His approach to polysemy is based upon this theory: a word has its core meaning, namely prototypical meaning, and although senses of the word are
slightly different from each other, they are related mutually through the core, and can be represented by a single word. In this way, multiple meanings of a word form a network and construct a complex category. Within this category, meanings extend from the core meaning and they are presumed to form a radial construction, and for this reason, Lakoff refers to this construction as radial category. Below is what Lakoff states about radial category, using mother as an example.

The category mother, …is structured radially with respect to a number of its subcategories: there is a central subcategory, defined by a cluster of converging cognitive models (the birth model, the nurturance model, etc.); in addition, there are noncentral extensions which are not specialized instances of the central subcategory, but rather are variants of it (adoptive mother, birth mother, foster mother, surrogate mother, etc.). These variants are not generated from the central model by general rules; instead, they are extended by convention and must be learned one by one. But the extensions are by no means random. The central model determines the possibilities for extension, together with the possible relations between the central model and the extension models. We will describe the extensions of a central model as being motivated by the central model plus certain general principles of extension (Lakoff, 1987: 91).

What is important here is ‘motivation’. Extensions from prototypical meaning have to be motivated. These motivations can be considered as sources of
polysemy in the framework of Lakoff’s study. There are three important sources of polysemy.

The first source is metaphor. Metaphor is a process of “mapping of the logic of one domain (usually, but not always, a concrete domain) onto another (usually more abstract) domain” (Taylor, 1989: 138). When a word is used in a different domain from its original domain through mapping, and the usage is well conventionalized, the word acquires a new sense. For example, the word foundation can be used to mean ‘basic idea’ as in expressions like to be without foundation. This is because an expression of a domain: BUILDING, is used to express concepts of another domain: THEORY.

The second source is metonymy. Metonymy is a cognitive process of “using one entity to refer to another that is related to it” (Lakoff and Johnson, 1980: 35). If a waitress says, The ham sandwich is waiting for his check, the expression the ham sandwich is used to refer to an actual person, the person who ordered a ham sandwich. Similarly, when we say, The kettle’s boiling, the object which is boiling is not the kettle itself, but ‘the water in the kettle’. As this usage is well conventionalized, the word kettle actually acquired the meaning ‘the water inside a kettle’.

The third source is image-schema transformation. Image schemas are cognitive structures, which “are directly derived from every day bodily experience” (Ungere and Schmid, 1996: 108). Lakoff asserts the importance of
image schemas for polysemy; “There are certain very natural relationships among image-schemas, and these motivate polysemy…” (Lakoff, 1987: 440). For instance, we use the preposition in in various situations such as The present is in the box and My friend is in the classroom. We, however, have a certain image of ‘relationship of something and its container’. By applying this image, we understand the meanings of expressions such as be in high school, be in love and be in trouble.

The cognitive linguists’ approach to polysemy is marked by its comprehensive nature. In the meanings extended from a core meaning, some meanings are still close to the core and have very much in common with the central meaning, while others extend too far and make it difficult to think that the meaning is derived from the core. In the framework of categorization, however, all stages of extended meanings can be explained by viewing the whole category of a word and not just focusing on individual phenomena.

2.4. Generative Lexicon approach

The ‘Generative Lexicon’ is Pustejovsky’s (1995) approach to the problem of lexical ambiguity, to the multiplicity of word meaning and to the question of how we are able to give an infinite number of senses to words using finite means. The main thesis of this approach is that a core set of word senses is used to generate a larger set of word senses when individual lexical items are combined with others in phrases and clauses. This system has four levels
(argument structure, event structure, qualia structure and lexical inheritance structure) which are connected by generative devices (type coercion, selective binding and co-composition) that provide the compositional interpretation of words in context.

Pustejovsky argues that former approaches to natural language semantics have ignored either the problem of how words are used in novel contexts or the creation of such new senses on the basis of compositionality. In language, words can have more than one meaning, but the means in which this extension of meaning is carried out can vary. Based on Weinreich’s (1963, 1964) two types of ambiguity, Pustejovsky distinguishes between contrastive and complementary ambiguity.

Pustejovsky’s Generative Lexicon proposes a model that addresses the question – neglected by Cognitive Semantics – of how senses are created. It states that a core set of word senses is used to generate a larger set of word senses when individual lexical items are combined with others in phrases and clauses. Although Pustejovsky is mainly concerned with non-metaphorical meanings, it seems an appropriate model to account for the way in which polysemy is created. In Chapter 4, I apply this model to the analysis of the polysemous senses in perception verbs.

In short, in Pustejovsky’s Generative Lexicon approach, polysemous senses are understood as manifestations of the same basic meaning in different
contexts. A strong compositionality model, consisting of four levels of representation for a lexical item, and generative connecting devices explains these senses. This framework seems the most suitable for explaining how the semantic content of different lexical items interacts in order to create polysemous senses.

2.5. Nunberg’s Pragmatic theory of polysemy

Nunberg (1979) adopts in his writings a narrower concept of polysemy. He does not discuss cases of non-systematic polysemy at all, but identifies polysemy simply with systematic polysemy. He tries to provide a theory that is capable of dealing with all instances of systematic polysemy, but he concentrates on cases of what Deane call closed referential polysemy. The central example of Nunberg is like the following:

(1) The ham sandwich is sitting at table 20.

His basic idea can be summarized in the following way: phenomena of systematic polysemy are phenomena of referential variation (as opposed to a variation of lexical meaning), and they are deducible from pragmatic factors. Nunberg (1979) is the first in a series of papers Nunberg devoted to this topic. He presents the problem of polysemy as a problem of reference, which he believes to be an essentially pragmatic issue. Reference depends on the co-operation between a speaker and a hearer, and can be considered successful if the hearer is able to identify what the speaker has in mind. One way of referring is by direct ostension or by direct linguistic specification of the intended referent. But another common
way of reference, which Nunberg claims to be the basis of (systematic) polysemy, is achieved in the way that the speaker does not directly refer to his intended referent, but instead to another entity which stands in a specific and obvious relationship with the referent. So, for example, one can point to a part of something to refer to the whole. This kind of reference is called deferred reference.

2.6. Kilgarriff’s (1992) approach

In his Ph.D. dissertation on polysemy, Kilgarriff follows a course that is quite characteristic of NLP (natural language processing) research. His approach is inspired by modern lexicography in general and computational lexicography in particular, corpus linguistics, computationally oriented linguistic theories like HPSG (Head-driven Phrase Structure Grammar) and issues of artificial intelligence research (most importantly, knowledge representation). Like Deane’s, his dissertation also includes an interesting and useful summary of literature, and it complements the former very well. Whereas Deane (1987) discusses the opinions on and findings about polysemy in traditional descriptive and cognitive linguistics, Kilgarriff seems to be unaware of most of this literature, but he presents a comprehensive picture of considerations about the topic in computational linguistics and lexicography up to 1992. After this summary, Kilgarriff presents in the rest of the first part of his dissertation something that can be considered a detailed reconstruction of exactly how a modern lexicographer goes about constructing the dictionary entry of a polysemous word on the basis of
a large corpus. This is certainly a valuable part of the dissertation, but contains few points that are really surprising or illuminating. Kilgarriff makes some extremely dubious assertions throughout this part, e.g. that lexicography is without doubt an integral part of linguistics or that lexicographers are the most skilled in deciding if two uses of a word instantiate altogether different words, different senses of the same word, or different interpretations within the same sense, and that, therefore, one can safely base lexical semantic analyses on dictionary entries.

Kilgarriff also notes that polysemy is no phenomenon the definite borders of which could be clearly defined. He (1992: 71-81) suggests that what is usually referred to as polysemy is actually a complex and gradual phenomenon at the crossroads of homonymy, collocation, analogy and alternation (so he actually chooses a two-dimensional one). Homonymy is understood in the usual way. Alternation is approximately equivalent to Deane’s (1998) open referential polysemy, i.e. the kind of meaning variation that is both systematic and is intimately connected to the structure of encyclopedic knowledge about the thing denoted by the word. Collocation is a kind of contextual meaning variation that is closely associated with the word one is interested in co-occurring with some particular other word, e.g. frontal meaning ‘direct and obvious’ “seems to occur only with attack and assault”. Analogy is another kind of contextual meaning variation that is quite interesting. Analogy is somewhat systematic but connected
neither to the knowledge representation underlying the use of y (like open polysemy), nor to general discourse context (like closed polysemy). This phenomenon has received very little attention in the literature and we know correspondingly little about the mechanisms that underlie it.

2.7. Systematic versus Non-systematic polysemy

The literature generally assumes at least since Apresjan, 1973 that there are two kinds of polysemy: regular (or systematic) polysemy and irregular (or non-systematic) polysemy. Systematic polysemy is that kind of polysemy where the relation between the interpretations $a_1$ and $a_2$ of a word $A$ is the same as between the interpretations $b_1$ and $b_2$ of a word $B$, and there are parallel sets of meanings for several further words as well.

The literature seems to be most divided (at least in terms of terminology) with regard to these two kinds of polysemy. Most work from the 90s. calls just systematic polysemy polysemy and considers non-systematic polysemy instances of homonymy and therefore theoretically uninteresting. On the other hand, most of the early texts on polysemy from before the 80s took little notice of systematic polysemy and therefore called non-systematic polysemy polysemy. And finally, there are several authors who treat polysemy as more general common term encompassing both kinds of polysemy (referring to kinds of polysemy by adjectival specification or by introducing new terms).
2.8. Deane’s theory of polysemy

Paul Deane (1987) was very thoroughly collected, reviewed, and summarized the entire literature up to 1987 in his Ph.D. dissertation. His approach is definitely in the line with the general holistic treatment of polysemy insofar as he also claims polysemy to derive from general cognitive strategies. More specifically, he outlines a theory “in which word meaning is closely integrated with certain kinds of extra-linguistic knowledge (i.e., the cultural background, expressed in the form of ICMs (Idealized Cognitive Model)), without such information, it would be impossible to account consistently for polysemy” (emphasis added-G.P.). However, contrary to Geeraerts (1993) and Tuggy (1993), he maintains that it is reasonable to distinguish polysemy from both homonymy and generality, and also to treat systematic and non-systematic phenomena, which he considers to be instances of two kinds of polysemy, separately.

2.9. Principled polysemy approach (Evans 2005)

Evans (2005) employs the principled polysemy approach to lexical concepts arguing that the lexeme time constitutes a lexical category of distinct senses instantiated in semantic memory. The notion of ‘lexical concept’ is used interchangeably with sense, which is the central theoretical construct of his theory. Principled polysemy is an approach which seeks to account for the meanings associated with words as not being absolute and fixed, but rather as being capable
of changing over time. Hence, in this qualified sense lexical concepts are treated as being mutable and dynamic in nature.

2.10. **Natural Semantic Meta-language (NSM)**

This approach is originated by Anna Wierzbicka (1996) later used by Cliff Goddard (1998) and Goddard & Wierzbicka (2002). Goddard uses this framework of Natural Semantic Meta-language to develop and justify semantic explications for the common ordinary meanings of the polysemous word *Culture* in the article ‘The lexical semantics of *culture*’. He proposes a set of semantic explications framed in terms of empirically established universal semantic primes such as PEOPLE, THINK, DO, LIVE, NOT, LIKE, THE SAME, AND OTHER.

2.11. **Relational approach (Fellbaum 1990)**

In relational models of the lexicon, words are organized according to their meanings using semantic relations or links to form a semantic network. Like prototypical models, the relational model works with semantic domains. In addition, it ‘attempts to make explicit the structural organization that is implicit in other models, and describes how the elements of a domain are related to each other’ (Evens 1988). Ideally, in a relational model of the lexicon, knowing the meaning of a word is knowing the word’s location in the semantic space of the lexicon.
2.12. The lexicographers

Lexicographers have generally written remarkably little about the nature of multiple word senses. On the one hand, this is surprising, given that it is an issue a lexicographer must struggle with for most of the dictionary entries they ever write. On the other, lexicographers are by and large concerned with writing dictionaries, not theorizing about writing dictionaries. What has been written about the task of lexicography is mostly concerned with more practical matters like how to select and alphabeticise the headword list or collocations, and how to represent syntactic, phonological, or usage information. While stating that the specification of word meaning is the central task for the lexicographer (Zgusta, 1971, p 23) and the division of a word’s meanings into senses is a central part of that, Zgusta’s authoritative handbook gives little guidance beyond admonishments to avoid making too many, or too few, distinctions (pp 66-67).

One reason is this. Theoretical studies of dictionaries and lexicography inevitably fall within the domain of linguistics. But linguistics has proved itself of very limited use to the lexicographer. Zgusta warns him or her “not to be too impressed by the basic uncertainty concerning the nature of lexical meaning” (p 24) and, until recently, the methodology of researchers in linguistics has almost always been too contentious and the conclusions insufficiently specific to be of practical use. More recently, particularly since Hornby’s Oxford Advanced Learners Dictionary (first published 1948), lexicography has been increasingly
informed by the results of work on syntax in linguistics, but word meaning has not been a particularly fruitful area of linguistic research. So there the situation remains the same. Lexical semantics research can learn from dictionaries, the lexicographer may conclude, but there is little about word meaning that lexicographers can learn from linguistics. If lexicographers were to stray into the theory of word sense distinctions, they would be straying into notoriously treacherous territory; better to stick to the practice.

Two articles in the lexicographic literature directly addressing the variety of senses a word may have are Malakhovski (1987) and Robins (1987). The former presents a taxonomy which distinguishes, amongst sense pairs which are not purely coincidental, distinctions in syntax alone, distinctions in meaning alone, and distinctions in both. But after declaring the hperlexeme (which includes all those senses which vary from each other in either syntax alone or in a more complex pattern of syntax and meaning) the ‘main structural unit of modern English vocabulary’ (p 48) the analysis goes no deeper. Robins’s article is a discussion of the impossibility of a rigid distinction between homonymy and polysemy. His quest for principles for distinguishing the two is not successful.

A more rewarding pair of articles is Ayto (1983) and Stock (1983). Ayto asks:

*What set of procedures do lexicographers have available to them to pin down those protean entities, ‘meanings’? ... How do they decide what, for the purposes of a dictionary, constitutes the meaning of a word, and*
where, in the case of polysemous words, one meaning ends and the next begins? (p 89).