

**Coreness and Frequency: Concepts
and Applications
to Vocabulary Teaching and
Learning**

**مفهومى الـ (Coreness and Frequency)
وتطبيقاتهما في تدريس وتعليم المعاني**

**Asst. Lect.
Ihssan A. J. Al-Muslimawi
University of Kufa- College of Arts**

Coreness and Frequency.....	(96)
------------------------------------	--------------

Coreness and Frequency: Concepts and Applications

to Vocabulary Teaching and Learning

مفهومي الـ (Coreness and Frequency)

وتطبيقاتهما في تدريس وتعليم المعاني

Asst. Lect.

Ihssan A. J. Al-Muslimawi

University of Kufa- College of Arts

Abstract

The notions of 'Coreness and Frequency' in English are becoming increasingly prominent issues in applied linguistics. Numerous studies have argued that both concepts are not new topics or controversial issues, because they are applicable to various analyses and definitions. Several studies (such as: Carter, 1998; Durkin, 1986; Lee, 2001; McCarthy, 1999) have revealed the core vocabulary characteristics of specific words and suggested parameters for any word to qualify as a core word in practical texts. Others (such as: Nation, 2001; Schmitt, 2000; Schmitt & McCarthy, 2002) have graded vocabulary and created word-lists targeting to various levels for testing, teaching, and pedagogical purposes.

Key words: Coreness, Core Vocabulary (CV), Frequency.

Introduction

In English, the concepts of coreness and frequency would seem to be intuitively uncontroversial and plausible. The idea of the existence of a general, lexical item or central core is taken almost for granted for many language teachers and linguists. For core vocabulary, indeed, there is no universally approved, ready-made list of "the core words of the English language" to be existed, because of the multiple definitions and interpretations amenable to this concept. For one thing would it be useful to specify one set of core words for one component of the language "spoken language", and another set for the other "written language"; in doing so, it

seems to act against the very notion of “coreness” as something that cuts across all genres in the language as a whole. Although it is famous that the vocabulary of spoken language is quite different from that of written language, for instance, it might be debated that these differentiations are noncore by definitions because core lexis is precisely that which is somehow central to the language as a whole and thus not particular to any register.

Therefore, when the notion of “coreness” is conducted more carefully, it becomes clear that many various definitions or approaches can be taken. In addition, it could be argued that “there are several core vocabularies of core vocabularies rather than a completely unitary and discrete core vocabulary” (Carter, 1987, p.33). In this article, some of the possible definitions of core vocabulary will be presented. This suggests that the many different possible concepts of core vocabulary highlight the essential ambiguity of the notion, i.e., the concept of core vocabulary has multifaceted nature.

As for frequency, Nation (2001) indicates that studies of native speakers’ vocabulary seem to suggest that second language learners need to know very large numbers of words. While this may be useful in the long term, it is not an essential short term goal. However, frequency based studies state that this is not so, because there are some words much more useful than others. For example, one might distinguish in a particular text many kinds of vocabulary such as high frequency words, academic words, technical words, low frequency words, and so on. Because the learner of the language needs to acquire both the variety of meanings and the form of a given lexical item, then learning vocabulary is not an easy process. For beginner learners the main question, of course, is ‘where to start’. General vocabulary wordlists have used frequency measures to aid in this process by providing common vocabulary items that happen frequently across different texts

The first and second sections of this paper address these two concepts: coreness and frequency respectively, as they cling together to the teaching and learning of vocabulary for second

language learners. The final section sheds some light on how these two notions will be applied to vocabulary teaching and learning.

The present study aims at contributing to the understanding of the coreness and frequency concepts. A fundamental aim is to discuss and provide a number of definitions for determining the more core vocabulary than the less core vocabulary and identifying which ones are more common and which ones are marginal. The work also suggests some methods to the applications of the two concepts.

1- Core Vocabulary (CV)

The notion of ‘coreness’ would be best defined through the different tests, which are measured to assess the more significant core words. To date, different methods have been introduced and taken to measure a core vocabulary. Carter (1998) argues that “there are many several core vocabularies rather than a completely unitary and discrete core vocabulary”. That is, the core vocabulary can be seen as one with a universal meaning that some words are more basic than others. Obviously, I should explore some possible concepts of core vocabulary as the ‘coreness’ is characterized by its multifaceted nature. As it is conducted by many studies (such as: Carter, 1998, p.36; Carter, 1987, p.179; Durkin, 1986, p.63; Lee, 2001, p.251-2), there are two main kinds of tests:

1-1: The Syntactic and Semantic Relations between Core Words

This shows how core vocabulary could be useful to be incorporated in building a well-designed and systematic semantic description of the language system (Carter, 1998). The next series of tests use syntactic and semantic relations among words. The fundamental concept underlying them is that core words are generic rather than specific.

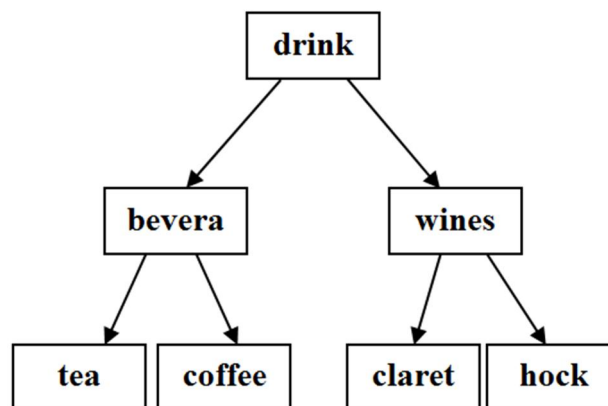
- Nuclear words can be replaced with non-nuclear, but not the vice-versa. For example, the words dine, devour, gobble, eat, stuff, gormandize all of them could be semantically defined as eat, but it would be not meaningful to define eat by the words in the set. In this case, eat is the more core word (Carter, 1998).

Coreness and Frequency.....(100)

- The more core a word is, the more obvious antonyms it has, i-e, it is easy to find an antonym to nuclear word. For example, the words good, bad; clean, dirty have predictable antonyms. However, it is so difficult to locate an antonym to words like excellence, spotless (Durkin, 1986,p.70-1).
- The collocational properties of the words provide with another test to core vocabulary. Collocation is the study of the words and their 'company' (Hunston& Francis,1999,p.230). Some lexical words co-occur with each other more or less strongly in the discourse. Accordingly, the core word has a wide collocational range in metaphorical use. In the following example, the feature (+) indicates that the word on the left collocates with the words at the top. In this case, 'dodge' is the more core word, because it collocates with most words or expressions.¹

	A blow	The police	The issue	Military service	One's duty
dodge	+	+	+	+	
evade			+	+	
duck			+		
shirk					+

- In his book, Durkin (1986) underlines that the core vocabulary has the property of extension. This test has a close relation with the latter since someone should consider the number of entries that a word (lexeme) has with different meanings. For example, Collins English Dictionary lists the following figures:
 - run (83), sprint (3).
 - house (28), mansion (5), villa (3).
 - give (29), award (4), denote (1).It is clear, therefore, that the word with different readings specified with different key words is showing more coreness than others.
- The vocabulary of a language is hierarchically structured, i-e, more general lexemes come at the top, and more specific words are sub-divisions as in the following instance:



So, 'drink' is a superordinate, i-e, beverages are kinds of drinks, which have the relation of inclusion 'hyponymy' (Aitchison, 1999, p.89). Consequently, the more core word is characterized by the common inclusion under a 'generic'.

1-2: Pragmatic Neutrality of Core Vocabulary.

In his article, Carter (1987) indicates that some words are more discursively neutral than others; they have the widespread usage in pragmatic contexts as unmarked or unexpressive. In other words, core words are pragmatically neutral in the sense that they convey no information about the situation of utterance. The following tests have to do with the relationship between words and social context.

- Culture-free: the more core a word is, the less probably it is to be restricted to culture-specific usage (Carter, 1998, p.41). Durkin (1986) argues that this notion should be relative to neutrality of specific geographical areas. He discusses, for example, that it would be neutral to the dimensions of human shape communicated by the word 'thin' in most Western cultures; however, it would be expressed for particular African cultures.
- 'Summary' is another genre to highlight core vocabulary. Informants are, as example, asked to summarize Hemingway's short story "Cat in the Rain", they prefer to use the word 'cat' in their summary. 'Cat' appears (13) while 'kitty' (6) times, although 'kitty' is a popular in the original text (Durkin, 2001, p.66). This leads us to infer that the core vocabulary is more preferable than non-core in summarizing the original texts.

Coreness and Frequency.....(102)

- The core vocabulary does not carry marked connotation or association. The connotational neutrality of nuclear words could be measured by analyzing our stylistic merits, with the hypothesis being that the core vocabulary tends to assume the value 0 on the scales (Stede, 1993,p.456).
- The core vocabulary does not state the genre of discourse from which it belongs. For instance, we know the words port, and starboard are related to navigation or nautical contexts while left, and right are not (Carter, 1998,p.43).
- The last point to be considered is that the core vocabulary is neutral in respect to tenor of discourse, i-e, it is not restricted to formal, causal, or slang uses. For instance, besides the word 'drunk' we have the formal words, inebriated and intoxicated and a number of colloquial, smashed, sozzled as Durkin (2001) argues.

1- Frequency

The 'frequency' notion has been adopted by many researchers (among them: Nation, 1997; Laufer, 1997; Carter, 1998; Schmitt, 2000; Nation, 2006). This reflects its importance to show to language learners which words are most frequent and useful for learning and teaching of vocabulary. To provide awareness about the frequency concept, researchers and learners should be acquainted with several factors:

- 1- **Corpus:** It is a collection of naturally occurring language to be analyzed linguistically. Corpora are the major source for counting word frequency or how a specific word occurs frequently. Collections of these corpora have been developed to be computerized, such as COBUILD, CIC, and BNC corpora, after they were selected manually, as the 'BUC and LOB corpora, depending on how many words a person is exposed to in real life and which areas are included (Schmitt, 2000,p.68-9).
- 2- We should consider whether the corpus was written, spoken or both as they differ significantly (McCarthy, 1990,p.67). For example, the occurrences of nouns in spoken language are more frequent than written tokens. Large corpora have been taken

from written discourse because it is easy to collect; conversely, the content words happen in spoken more than written discourse such as got, well, right(Schmitt, 2000,p.73-4)

- 3- Frequency and Range:The words should occur in a broad range of texts even if they occur in some form or other in most texts groupings (Schmitt & McCarthy, 2002,p.18). Frequent words should also represent language as a whole, regarding all subject areas from a wide range of genres (Carter, 1998,p.235).
- 4- Word Families: Attention should be paid to the forms and uses as being members of the word family. For instance, should governor be regarded as part of govern word family? (Schmitt & McCarthy, 2002,p.18). Nation (2001) states that word count should be checked for lemmatization as a word family includes a headword, its derived and inflected forms.
- 5- There are also many other criteria should be considered in the frequency notion such as idioms and expressions which occur as multi-word items frequently like never mind, good morning. The variations of meaning, collocations and uses of the word , and the restrictions on the word use, are all significant issues for deciding what goes into the most frequent word list(Schmitt & McCarthy, 2002,p.18-19)

Examples of the word frequency lists are numerous. The classic list is the 2,000 headword of West's General Service List(1953). Its frequency is based upon 5 million word written corpus(Schmitt & McCarthy, 2002, p.13-14). Currently, the most noticeably is the New Academic World List, which was developed by Coxhead for her MA thesis(Thomas,2010). It contains 570 word families,which was composed from 3.5 million written corpus. According to range, frequency, and uniformity principles, the AWL has been selected. For range, it had to occur across disciplines like Arts, Commerce, Science, and Law sections and seven subgroups for each which ensures its usefulness for all learners. For frequency, AWL families had to occur over 100 times in 3.5 million words Academic Corpusto ensure that the words will be met in a reasonable number of academic texts. It should also occur over 10 times in each faculty of the academic texts(Coxhead, 2000,p.213). Narrow range words,

proper nouns, Latin forms, and words including in the first 2,000 words of English are all excluded from the AWL.

3- The Applications of Coreness and Frequency Notions to Vocabulary Teaching and Learning

One of the areas that needs greater attention and requires a constant reference is how to apply the coreness and frequency notions to language teaching and learning of vocabulary. Generally, there is no specific way for teaching and learning these two ideas; it depends on many factors, which include the level of the students, the words that are target, the school system and the syllabus design (Schmitt, 2000, p.142).

Teachers need to develop awareness of what words to teach first and to which level of education. In this regard, it is still the best way for the second language (L2) learners is the 2000 most frequent words of English as many authors (Among them: Nation, 2001, p.15-16; Schmitt & McCarthy, 2002, p.9-10; Schmitt, 2000, p.142-3; Carter, 1998, p.206-7; Coxhead, 2000, p.227-8) agree that learners who know the 2000 words should be able to understand almost 80% of the words in text. Therefore, a vocabulary of about 2000 words would be a realistic goal and learners who know most high-frequency words help teachers to make decisions about what to teach and in what order. For instance, we can see many idioms that are rare, so we can teach them later. On the other hand, we can see a large amount of vocabulary students talk about most and teach them first, leaving the low-frequency words until later.

Because of the size and variety of the vocabulary kinds, it is challenging for learners to learn vocabulary, containing single words, collocations, phrases 'chunks', and strategic vocabulary, besides structural patterning, fixed expression, and idioms, that is, it would be tedious lessons to teach everything about the word. Obviously, Nation (2001) suggests that the most frequent words can be taught through direct teaching and planned meeting with the words, as well as he clarifies that teachers need to make choices how much they teach on first presentation of the vocabulary. Teachers should use level tests or interview learners to determine

whether the focus on a high, intermediate, or pre-intermediate level, whether they first teach CV, high-frequency, low-frequency, or academic vocabulary. For example, if the goal is to enrich vocabulary knowledge, 20 words a week might be enough, but if the goal of teaching is to add new vocabulary, then the number should be larger (Schmitt & Schmitt, 1995, p.140)

The choices we make are subject to factors of frequency, usefulness for the classroom and how the word is easy to learn and teach. For example, we can express to learners how to know similar and opposite meanings for a set of vocabulary we are teaching, how to look at various meanings of the same item as incremental process, and how to return to a word they have already learned to add more information about it (Schmitt, 2000, p.157-8).

As for learning, first, materials need to be practiced and presented in natural contexts the vocabulary that is current, suitable for learners' needs, and frequent. Secondly, syllabus should help students become better learners of vocabulary by teaching several strategies and techniques, not only inside the classroom, but also outside the classroom.

There are many principles that help students learn vocabulary. One of the important classroom learning strategies is to teach students how to ask for words that they do not know, e.g. 'what does----mean?'. Another strategy is to help students get into the habit of 'noticing', what each item is, e.g. 'a single word, a phrase, its form, a collocation, etc.', and for what purpose, e.g. 'active use or passive recognition', (Nation, 2001, p.62-4). Teachers' different methods and a variety of presenting the vocabulary lists are other attractive techniques to enhance students knowledge of high-frequency words; for instance, using pictures, stories text types, web pages, conversations, etc., these will cater for the different needs and interests of students (Schmitt & McCarthy, 2002, p.244-5). Furthermore, retrieving, direct and incidental learning, guessing, and using games all seem to be effective learning styles to apply both notions.²

Conclusion

As seen earlier, literature indicates that there is no specific direct definition to the notions of coreness and frequency. Different tests, factors, and parameters should be considered to define each concept. Similarly, applying these two ideas to vocabulary teaching and learning are subject to the way, goal, and level of education. Teachers can start with the most core and frequent vocabulary, learnable, and useful to their students. They can make the syllabus materials more effective through using varied teaching and learning contexts.

مفهومَي الـ (Coreness and Frequency)

وتطبيقاتهما في تدريس وتعليم المعاني

ملخص البحث

قدم هذا البحث مفهومين أساسيين من مفاهيم علم اللغة التطبيقي، هما (Coreness) و (Frequency) بإعتبارهما موضوعين مهمين في تدريس وتعليم المعاني لتعليمي اللغة الانكليزية. إذ بين البحث اهم التعاريف والمفاهيم المتعلقة بهما وأشار الى أن الدراسات السابقة قد إستندت الى عدة إختبارات وتحاليل لكي تعرفهما. وإنقسم البحث الى ثلاثة أقسام، ففي قسمه الاول بين مفهوم (Coreness) وكشف عن اهم المميزات والخصائص التي من خلالها يمكن لنا معرفة المعاني التي تتسم بالـ (Coreness). وفي القسم الثاني منه، وضح البحث مفهوم (Frequency) وهو تردد عدد مرات الكلمة واهم المعايير التي يجب معرفتها في هذا التردد. أما في القسم الاخير من البحث، فقد تناول كيفية تطبيق هذين المفهومين في تدريس وتعليم المعاني لتعليمي اللغة الانكليزية.

References

- Aitchison, J., (1999). Linguistics. London: Hodder Headline.

Coreness and Frequency.....(107)

- Carter, R. (1987). Is there a core vocabulary? Some Implications for Language Teaching. 'Applied Linguistics' Journal vol. (8), 178-193.
- Carter, R. (1998). Applied Linguistics Perspectives (2nded.). London: Routledge.
- Doughty, C. & Williams, J. (Eds.). (1998). Focus on Form in Classroom Second Language Acquisition. Cambridge: CUP.
- Durkin, K. (1986). Language Development in the School Years. Croom Helm. [Accessed via Google Scholar]. Retrieved from <http://www.uni-trier.de/fileadmin/fb2/ANG/Linguistik/Stubbs/stubbs-1986-nuclear-vocab.pdf>
- Hunston, S. & Francis, G. (1999). Pattern Grammar: A corpus-driven approach to the lexical grammar of English, (Studies in Corpus Linguistics). Amsterdam: John Benjamins.
- Laufer, B. (1997). What's in a word that makes it hard or easy? Some intralexical factors that affect the learning of words.", Vocabulary: Description, Acquisition and Pedagogy, Cambridge: Cambridge University Press, pp. 140–155.
- Lee, D. Y. W. (2001). Defining core vocabulary and tracking its distribution across spoken and written genres: Evidence of a gradience of variation from the British National Corpus. Journal of English Linguistics, vol. 29(3), 250-287.
- McCarthy, M. (1990). Vocabulary.Oxford: CUP. [Accessible via Google Scholar].
- McCarthy, M. (1999). What constitutes a basic vocabulary for spoken communication? Studies in English Language and Literature Journal, vol. (1), 233-249.
- Nation, I.S.P. (1997). "Vocabulary size, text coverage, and word lists", in Schmitt; McCarthy, Vocabulary: Description, Acquisition and Pedagogy. Cambridge: Cambridge University Press, pp. 6–19.
- Nation, I. S. P. (2001). Learning Vocabulary in Another Language. Cambridge: CUP.
- Nation, I.S.P. (2006). Language Education - Vocabulary. Oxford: Encyclopedia of Language & Linguistics, pp. 494–499.
- Schmitt, N. & Schmitt, D. (1995). Vocabulary Notebooks: theoretical underpinnings and practical suggestions.ELT Journal, 49(2): 133-43.

Coreness and Frequency.....(108)

- Schmitt, N. & McCarthy, M. (Ed.). (1997). Vocabulary Description, Acquisition and Pedagogy. Cambridge: CUP.
- Schmitt, N. (2000). Vocabulary in Language Teaching. Cambridge: CUP.
- Stede, M. (1993). Lexical Choice Criteria in Language Generation. [Online] Available at: <http://ucrel.lancs.ac.uk/acl/E/E93/E93-1055.pdf> [Accessed 22 October 2013].
- Wu, W. (n.d.). Lexical Collocation: One Way to Make Vocabulary Active. [Online article] Available at: <http://people.chu.edu.tw/~wswu/publications/papers/conferences/01.pdf> [Accessed 19 October 2013].