LEIDEN UNIVERSITY

FACULTY OF HUMANITIES

LEIDEN UNIVERSITY CENTER FOR LINGUISTICS

Universiteit Leiden

Graduation Thesis

Relative Frequency Effects in English & Modern Greek: An Experimental Study

Antonis Zolotas

s1917218

Leiden, 2018

LEIDEN UNIVERSITY

FACULTY OF HUMANITIES

LEIDEN UNIVERSITY CENTER FOR LINGUISTICS

Graduation Thesis

Relative Frequency Effects in English & Modern Greek: An Experimental Study

Antonis Zolotas

s1917218

Supervisor

Dr. E.D.Botma

Leiden, 2018

Copyrights

I would like to thank the following people, whose help has been invaluable:

Dr. *E.D. Botma* for his guidance, patience and useful advices during the creation of this thesis.

My whole *family* for supporting me all the time and especially during my Masters.

My flatmate *Barry* for making my stay in the Netherlands much more enjoyable.

My friend *Marios* for helping me with the frequency values.

Everyone that participated in the experiments I conducted for the purposes of this thesis.

ABSTRACT

The goal of this thesis is to determine the factors that affect the decomposability of affixed words in a cross linguistic level by conducting two experiments, the first in English and the second in Modern Greek language. With the term decomposability I refer to the property of affixed words to be broken down to their basic components in order to be semantically defined. Native speakers are able to distinguish and seperate their language's complex words into their basic morphemes. One of the main factors that is suspected to affect the decomposability of affixed words is the relative frequency of the words themselves. After I provide sufficient information about the specifics of relative frequency and other possible factors that can affect the decomposability of affixed words, I will present in detail the two experiments that were performed along with their results. The thesis will conclude with a discussion of the results before determining whether the examined factors share the same effects in both languages and examining if the possibility of other factors that affect the phenomenon of decomposability in a cross-linguistic level may exist or not.

TABLE OF CONTENTS

INTRODUCTION

This graduation thesis has been created within the context of the English Word Formation subject for the MA Linguistics: Theoretical and Experimental Linguistics. The study deals with a set of experiments that were conducted in two languages, Modern Greek and English, in order to examine the factors that affect the decomposability of affixed words. Here I will examine the *decomposability* phenomenon. i.e. the property of affixed words to be decomposed into their embedded components. Native speakers are able to distinguish and seperate their language's complex words. When trying to 'break' a word into its components, some factors determine the convenience of this procedure; some words are more easily decomposed as opposed to others. I will attempt to figure out these factors with a set of experiments and attempt to provide an answer to the question of whether these factors differ across languages or at the very least whether they affect the decomposability of affixed words in the same way in languages other than English. The thesis will be divided in six (6) chapters excluding the Introduction.

In the *first chapter* I will present the theoretical framework, on which the survey and the thesis are based on. It consists of a small introduction in morphology, followed by decomposability and relative frequency, while briefly reviewing the literature that is related to the connection of the two linguistic phenomena.

In the *second and the third chapter* I will provide information about the languages that are going to be examined during this thesis presenting each language's history and general linguistic characteristics as far as word formation is concerned. These characteristics could affect the factors with which a native speaker of a morphologically complex language decomposes a complex word in contrast to a non morphologically complex language.

The *fourth chapter* involves the methodology of the two experiments that were performed in Greek and English. Here the specifics about the words that were chosen along with the participants and the materials used will be given.

The *fifth chapter* will focus on the first experiment, performed for the English language. It is similar to the experiment that was conducted by Jennifer Hay (2001) for the purpose of proving the existence of a relative frequency effect. The experiment's goal for the purposes

of this thesis, however, is to determine which factors play a major role in decomposing affixed words.

The *sixth chapter* will examine the second experiment that was performed for the Modern Greek language and will set the basis of the discussion in the following section.

In the *seventh chapter* I will summarize both experiments and continue with a discussion of their results. These results will be the main point for conclusion on the cross-linguistic aspect of decomposability.

The *eighth and final chapter* will conclude the thesis with a recapitulation of what was examined and set new questions for future surveys that can be conducted on a bigger variety of languages for the decomposability phenomenon in linguistics.

CHAPTER 1: Decomposability and Relative Frequency

1.1 General on Morphological Background

Morphology is one of linguistics' core branches that is responsible for the study of words; analyzing how they are formed, their components' structure and their relationship with other words of the same language. Morphology has three major fields of studies, which are inflection, derivation and composition. The former is responsible for the creation of different forms of the same lexeme by making use of the inflectional affixes in a language. In other words it is the modification of a word in order to express different grammatical categories, such as the person, number and case but without changing the word's core meaning and it is expressed with the addition of at least one inflectional affix to the word. On the other hand, derivation and composition are processes that result in the creation of different lexemes, either by using derivational affixes to the word's stem or by combining entirely different words to form a new word. Derivation and composition are the two kinds of morphological word formation.

Word formation refers to the creation of new words. These words are formed by combining smaller elements to form larger units with a more complex meaning (Plag, 2003). These newly formed words are morphologically complex words and the units that these words consist of are called morphemes. Morphemes are a word's smallest meaningful units and they can be either free, which are the morphemes whose boundaries overlap a word's boundaries, or bound, which are the morphemes that must be combined with other morphemes in order to be a part of a word. The morphemes can be further categorized into stems and affixes. The stem is the base for the formation of a word and in certain languages like English a stem can also be a word at the same time. Affixes are combined with a stem in order to form a new lexeme or different forms of the same lexeme. They are further divided mainly into prefixes and suffixes depending on if they are put before or after the stem respectively.

At this point I must further elaborate on the term lexeme. A lexeme is an abstract unit of lexical meaning that represents a set of different forms of a single word. The lexeme is conventionally depicted with capital letters, e.g. *RUN*. The different forms *run*, *runs*, *ran* and *running* belong to the same lexeme *RUN*. This term holds great important in the field

of Morphology, since it can be used to differentiate other terms in linguistics. For example, inflection and derivation can be defined by the difference that the former's rules create forms of the same lexeme, while the latter's rules create a different lexeme altogether.

The focus of this thesis and the experiments included is on word formation rules and specifically on derivation, the word formation process occurring mainly by combining a word's stem with a derivational affix, assuming that the language has combinatorial properties. This process is called affixation. In the list below I mention some characteristic examples of affixation along with other word formation processes.

(1) a. work-er
b. in- activ-ity
c. sing – song
d. switch – switch
e. import – import

In (1a) there is the word/stem *work* and then the derivational suffix *–er* can be added. As a result the derivative *worker* will be formed. It is possible to use multiple affixes on a stem to form a word or even use an affix on an already formed derivative like in (1b). The derivative *activity*, consisted of the word/stem *active* and the suffix *–ity,* can form a new derivative, if the prefix *in-* is added, resulting in the form *inactivity*. Besides using affixes derivation can also occur with other processes like ablaut, the process of forming a derivative by changing the vowel in the word's stem, like in (1c) *sing – song*, and conversion, which is also called zero derivation, occurring when there are two words having the same form, (1d) *switch* (N) – *(to) switch* (V), but each one belongs in a different lexical category. One of the forms is a verb, whereas the other is a noun. The category, to which the form belongs, is dependent on its place in a sentence.

(2) a. The *switch* is off.
Noun

b. The kids *switch* on the TV.
Verb

This process is very common in languages such as English and the noun can convert into a verb and vice versa. There is also the process of changing the stress of a word, which is connected with the word's suprasegmental side (Ralli, 2005), a process similar to

conversion, albeit strictly speaking it is not the same. Thereby we have examples like (1e) impórt (V) – ímport (N). The verb has the stress on its final syllable, while the noun has the stress on the initial syllable.

As mentioned, the most common derivational process is affixation. Affixes are words' bound morphemes, meaning that they can't appear freely and they necessarily have to be attached to other morphemes in order to form words. We cannot use, for example, the affix **er* on its own during speech, but it can combined with a word/stem like *work* to form *worker*, an affixed word.

1.2 Decomposability of Affixed Words

Similar to words, tracing the boundaries of morphemes in certain words can prove to be a difficult task. These difficulties are mostly due to the interaction between Morphology and Phonology, as well as the interaction between languages which results in loanwords, words that are adopted from one language and incorporated into another without being translated.

(3) a. irresponsible

b. astronomic

In example 3a the prefix *in-* is combined with the word *responsible*, which in turn consists of the word *response* and the suffix *-able*. The sound *n* in the prefix assimilates with the sound *r* because of its existence in the following word *responsible*. This phonological process is called nasal assimilation. Example 3b depicts a loanword from Modern Greek, the word *astronomic*, which means something that is connected with astronomy or something that is extremely large. The word itself can be broken down in the morphemes *astronom-* and *-ic*. The issue with this example is that it is not certain whether we should further break down the morpheme *astronom-* into *astro-* and *nom-*. In Modern Greek the morpheme *astronom-* is a compound, formed by combining the word *astro* (=star) and the bound stem *-nom-* (=arranging, regulating). The bound stem *-nom-* is not an English morpheme, since it is not a meaningful unit in that language, instead it was adopted from Modern Greek and the average English native speaker is not expected to know *astronom-*'s etymology. Consequently, when decomposing the word *astronomic* the morpheme *astronom-* should remain undivided.

Decomposition is a process that, in principle at least, is a part of the decoding of a language on the part of listeners and readers. In the same way that we are able to form words by combining stems with affixes, we can also decompose them by breaking the words down to their constituents in an attempt to understand the meaning of the affixed word's components and the word itself. According to the principle of compositionality in the case of linguistics, we can combine lexemes with affixes that hold a meaning to form another lexeme with a more complex meaning than its original morphemes. Subsequently, we can break down a complex form down to its constituent parts in order to interpret its meaning.

(4) a. unfriend
b. hater

The words listed above are newly formed words that have appeared in dictionaries such as the Oxford English Dictionary and Merriam-Webster Dictionary in recent years due to an increased use of social media, and consequently the Internet. Even if someone has not encountered these words in the past, by decomposing them down to their constituents it is rather easy to understand their meaning. The example (4a) *unfriend* can be broken into *un-* and *friend*. The prefix *un-* is combined with verbs in order to express the notion of reversing an action, in this occasion, the process of adding someone to their friend list. *Unfriend* is the process during which someone removes a person from the said list. The word *hater* (4b) is a combination of the verb *hate* and the suffix *-er*, a suffix that is added to verbs in order to express the person or thing that does an action indicated by the root verb. In this example it means the person who greatly dislikes a specified person or thing.

I will be referring to this property of affixed words with the term decomposability from now on. All affixed words are decomposable, although the degree varies per word and even the speaker. By examining the following examples;

(5) a. claustro-phob-ic
b. in-activ-ity

We notice that both are affixed words, yet the example (5a) *claustrophobic* (= a person with a fear of being shut in a confined space) has a lower degree of decomposability compared to (5b) *inactivity*. Both of them are certainly decomposable, however a speaker who specializes or has extensive knowledge in Medicine or Latin can completely identify

all components of the lexeme *claustrophobic* (claustrum + phobia + -ic) with more ease than another speaker with limited or no knowledge in either fields. Every native speaker possesses the ability to decompose complex words to its basic components, although the question at hand would be the following; if we had a group of speakers, who belong in the same social group(s) and age and share the same education, what would be the factors that would determine which words are more decomposable than others?

1.3 Factors Under Examination

The most important factor that determines the decomposability of complex, affixed words when compared to other words seems to be the frequency of the words, i.e. how often the native speakers access these words in their mental lexicon to use them in speech and writing.

In the majority of models of morphological processing, access to morphologically complex words in the mental lexicon is achieved by two routes; the whole word and the decomposed route (Plag, 2003). The first route supports the notion that a complex word is stored in the lexicon as is, while the second route argues that we store the decomposed elements in the lexicon. According to these models, the frequency holds an important role in determining the 'resting activation of lexical items' (Hay, 2001 & Plag, 2003) and each route strengthens the representation of different aspects. The whole word route strengthens the whole word representation, while the decomposed route strengthens the decomposed morphemes and the decomposability of words. The usual assumption is that since the absolute frequency correlates with the resting activation, it will be the most important factor in determining, which representation will be strengthened with a given word.

It is necessary, however, that I make a reference to the absolute and relative frequency at this point. Hay (2001) is an attempt to demonstrate the existence of a relative frequency effect. The widely assumed opinion at the time was that the absolute frequency of a morphologically complex word is highly related with non-compositionality, i.e. the word cannot be easily decomposed.

On the other hand, Hay argues that the (non-)compositionality of a word is related to the relative frequency, according to the morphological models' predictions, rather than the

absolute frequency. But what exactly is relative frequency? The relative frequency effect is the ratio of the frequency of a complex, i.e. derived, word to the frequency of the base that it derives from. Relative frequency measure how frequent the derived word is compared to its base. Hay (2001) attempts to prove that if the derived form of a base is more frequent than the base itself then it would be more likely to undergo a semantic drift. Likewise, if the derived form is less frequent than the base then it would be more tightly constrained to the characteristics and meaning of the base.

In order to prove the existence of this effect, after selecting 34 pairs of words from the CELEX database, Hay (2001) conducts a simple experiment, in which she asks her participants to choose which word from these pairs is more complex according to them. Each pair's words are placed in one of the two different columns, which she names Columns A and B. The A members of each pair are more frequent than the bases they are derived from, whereas the B members are less frequent than their base. The term frequency refers to the amount of instances that a word has appeared in a selected database.

Hay's prediction is that the A members will be rated less complex than B members. For example the word from Column A *immortal* is formed by the prefix *in-* which turns into *im-* due to nasal assimilation, a phonological process in which a sound changes into a nearby sound, and the adjective *mortal*. The word *immortal* is more frequent in the CELEX database than its base *mortal* and it would be rated as more complex than the word from Column B *immoral*, formed by the prefix *in-* and the adjective *moral*.

The results of the experiment were in favour of her hypothesis, successfully proving that a relative frequency effect does, in fact, exist. This also means that the relative frequency effect will determine which representation will be strengthened when we attempt to access a complex word in the lexicon. If the base is more frequent than the derivative, the decomposed route is strengthened, since the base will also be accessed when trying to access the complex word. If the derivative is more frequent than the base, the whole word route is strengthened, because the derived word is more likely to have undergone a semantic drift and would be stored in the mental lexicon as a whole.

Furthermore, she examined the relationship between absolute and relative frequency, discovering that absolute frequency has some connection with relative frequency and subsequently with non-compositionality. However relative frequency seems to be more connected with it.

Another factor that could possibly influence the decomposability of affixed words is where exactly the affixes are attached to the word or stem. The affixes are divided in categories depending on their position concerning the stem.

(6) a. *de*-form
b. colour-*less*
c. picoline → pi-*pe*-coline (English)
d. ktb → k*ata*b (Arabic)
(to) write he wrote
e. Spielen → *ge*-spiel-*t* (German)
(to) play have played

The categories, that the affixes belong to, consist of the **prefixes**, placed before the word's stem as in example (6a) and **suffixes**, appearing after the stem (6b). There are also some other affixes used in rare occasions like **infixes**, that intervene between the phonemes of the stem itself (6c) and in English are found in technical terminologies like chemistry's vocabulary. The example in (6c) *picoline* refers to three different isomers, a term used in Chemistry. The -*pe*- infix signifies complete hydrogenation, i.e. to treat the isomers with hydrogen. Moreover, there are the **transfixes**, phonemes that are required to appear in morphemes or words in order to form a word (6d) and are found mainly in templatic languages e.g. Arabic (Ralli, 2005) and **circumfixes**, where we have discontinuous morphemes to form certain types in languages like German in the verb's perfect tenses (6e) (Ralli,2005).

1.4 Decomposability on a Cross-linguistic Level

The second question that this thesis will attempt to answer is if the factors that affect decomposability will have a similar effect cross-linguistically or if they will vary across languages. Supposing that relative frequency is one of these factors and it holds an important role in decomposing affixed words in English, is it certain that it will be of the same importance or even have a similar effect on other languages, even if these languages have a morphology fairly similar to English?

In order to answer the aforementioned questions, two experiments will be conducted on two different languages. The experiments will be similar to Hay's original experiment that she made for the relative frequency effect, however in this case the experiments will have a different goal, i.e. to determine whether the factors that differentiate these words from each other affect their decomposition. The experiments will have the participants to choose between a pair of words asking which one of the two words is according to their own judgement more complex. The words in these pairs will have a different relative frequency between them, while other pairs will focus on the prefix vs. suffix issue, or on a combination of factors. It is expected that the participants will choose the affixed word with the lower frequency as being more complex, which would mean that these words are easier decomposed than the affixed words with a higher frequency than their base. It should be noted once more that this does not exclude the other word of the pair from having the ability to be decomposed, it naturally means that it has a lower degree of decomposability compared to its paired word. Further details of the experiments will be discussed in the following chapters.

CHAPTER 2: Languages Under Examination - English

Before describing the set-up of the experiments, it is necessary that I make a discussion about the two languages under examination, English and Modern Greek, as far as morphology is concerned. These two languages share common traits among them, but also have different characteristics.

They were ultimately chosen because of their similarities, since they belong in the same family of languages, yet possess important differences due to their morphology that may hold an important role in decomposability. In the following sections I will provide information and general linguistic characteristics, specific to their morphology, regarding the two languages. I will begin this analysis with English.

2.1 History of English

English belongs to the Indo-European family of languages and more specifically in the Germanic languages subdivision. Originating from the Proto-Germanic language, English had a lot of periods of development. The language, known first as Old English during the Anglo-Saxon era, was first brought to Britain by Germanic invaders and settlers, displacing the previously spoken Celtic languages.

It was later developed in Middle English during the invasion of the Normans in the 11th century, when the language was heavily influenced by the North Germanic languages, which were spoken by the Scandinavians, and more importantly by the French language, from which many loanwords entered the language. Another important feature of this period is that the orthography system was first established in the language and it is still in use even to this day.

Starting from about the middle of the 15th century (Crystal, 2003) the language developed into the Early Modern English. In this phase many of the French loans from the past eras were consolidated and new classical loanwords from languages such as Latin and Ancient Greek entered the language along with words from other European languages like German and Dutch. It was during this period that pronunciation changes like the Great Vowel Shift, which affected the qualities of most long vowels, occurred in the language.

The latest recorded development period in the language was nearly the beginning of the 18th century, when the language entered the Modern English era, whose main structure and characteristics are similar in most respects to this day. The spelling, the punctuation and the grammar are very close to the ones that are used today. However, despite the similarities, it is not completely identical to the language that is used in the 21st century, as English is continuously evolving like every other non-extinct language in the world. Many words that were used at the time despite having the same form/spelling, they had completely different meanings (Crystal, 2003).

The language has been exported to many parts of the world through the British Colonisation and remains the dominant language of a great number of countries like the United Kingdom, the United States of America, Canada, Ireland etc. Being one of the most widespread languages in the world due to the United States influence and the Internet's extensive use, it has become the modern day's lingua franca (Crystal, 2003), i.e. the language that is systematically used between people from all over the world that do not share the same native language in order to communicate. Since it is in constant contact with other languages across the world due to its current status, it is not unexpected that English will continue to evolve over time.

2.2 English Morphology

It is considered that English has a rather simple morphology compared to other European languages like French, Russian and Greek (Tesitelova, 1992 & Tran KM, 2014). Bauer et al (2013) mention that English lacks certain morphological phenomena such as root and pattern morphology and in other phenomena the language is very poor, as is the case with infixation, reduplication, ablaut and other sorts of internal modification. This does not mean, though, that the field of Morphology is not an interesting topic of research in English nor that it is not worth to study it.

2.2.1 Inflection & Derivation

In the previous chapter I mentioned that the smallest meaningful units of a word are called morphemes. Morphemes are divided in bound and free, depending on if they can stand on

their own or not. Carstairs-McCarthy (2002) states that a salient characteristic of English, which differentiates the language from others, is that a large amount of complex words contain a free morpheme at their core/root (root henceforth), as long as the root is not borrowed from languages such as French or Latin.

(7) a. *help*-ful
b. leg-*ible*

In the example 7a above the complex word *helpful* consists of two morphemes, *help* and *-ful*. The morpheme *help* is a free morpheme, as it can stand on its own in other sentences, while the morpheme *-ful* is an affix, a bound morpheme that must be combined with other morphemes in order to appear in a word. But not every word in English has a free morpheme as its root, as in example 7b the word *legible* has the morphemes *leg-* and *-ible* as its constituents and both morphemes are bound. The morpheme *leg-* has been introduced to English from the Latin language and does not appear in English as a free morpheme.

It is not uncommon that an English word is at the same time a root; in other words a free morpheme, meaning they can occur on their own. *Dog* and *house* are some examples of English stems that are also words and such words are called monomorphemic. On the other hand, the morphemes that are non-root and are attached to roots to form new lexemes or variations of the same lexeme are the affixes and are bound morphemes. Among the many types of affixes mentioned in *§1.3*, the English language includes suffixes, prefixes and infixes. As it was previously mentioned, infixes are very rare in English and can be found in technical terms like *pi-pe-coline*, in which the infix *-pe-* is added to the word *picoline* or intervening a complex word, as is the case with the infix *-bloody-* in the word abso-*bloody*-lutely (Plag, 2003). In general, English has considerably more suffixes than prefixes (Carstairs-McArthy, 2002) and a very small amount of infixes.

Affixes are categorised in derivational and inflectional affixes depending on if they create new lexemes or different forms of the same lexeme. The inflectional affixes in English create different forms of the same lexeme and they are encountered in almost every lexical category, as shown in the following examples.

(8) a. house-*s*
b. talk-*ed*

c. green-*er*

d. soon-*er*

Inflection affixes can be attached to nouns (8a), verbs (8b), adjectives (8c) and adverbs (8d). An interesting fact is that English contains only suffixes as inflectional affixes. These suffixes carry information about the number (8a), the case (8b) and comparison (8c, 8d). The remaining information about the gender, the case, the number and tense are mostly given by the word order in a sentence.

Regarding the comparative suffix -*er*, one can argue about whether the affix is derivational instead of inflectional. Carstairs-McCarthy (2002) in one of his examples explains that the reason that this suffix and the superlative suffix -*est* are considered inflectional is because there are certain grammatical contexts, in which anything else other than the comparative or superlative form of the lexeme would be grammatically unacceptable (9b), even if it is semantically appropriate. This theory can potentially be supported, if inflection's basic difference from derivation is taken into consideration and that is that it creates forms of the same and not a different lexeme.

(9) a. The field is greener than that one.

b. *The field is green than that one.

Inflection holds a rather modest role as opposed to other languages like German (Carstairs-McArthy, 2002), who have a deeper inflectional morphology. The limited number of different forms that a lexeme has in each lexical category due to the inflectional suffixes reflects this statement. The lexeme HOME has two different forms *house* and *houses*, while the lexeme GREEN has three different forms, *green, greener* and *greenest.* Bauer (1983) comments that due to the language's relatively low amount of inflectional suffixes, the inflectional categories tend to be fully productive.

On the other hand, the derivational affixes are responsible for the creation of different lexemes in English. As with the inflectional suffixes they are encountered in almost every lexical category.

(10) a. *re*-do

b. bottom-*less*

c. brother-*hood*

d. good-*ness*

e. coward-*ly*

Derivational affixes can be combined with verbs (10a), prepositions (10b), nouns (10c) and adjectives (10d, 10e). Derivational affixes in English can be either a prefix (10a) or suffix (10b-10e). This means that in English the prefixes can only be derivational, whereas the suffixes can be either inflectional or derivational.

An important property of the derivational affixes is that, since they form a new lexeme, they can also change the lexical category of the word to which they attach. This is clearly shown in examples 10b, 10d and 10e. The suffix -*less* forms an adjective after it is combined with the preposition *bottom*, while the suffixes -*ness* and -*ly* form a noun and adverb respectively when they combine with *good* and *coward*. However, it is not mandatory that the lexical category is changed. In examples 10a and 10c the prefix *re-* and the suffix -*hood* maintain the same category that the initial word belongs to.

At this point I must make a short reference to the adverbial suffix -*ly*. There has been a lot of controversial discussion on this topic with the main argument being whether it should be considered a derivational or an inflectional affix. Plag (2003) states that since -*ly* is "for the most part syntactically triggered and obligatory", it can be identified as an inflectional affix. However, he mentions that there are some formations, in which there is a difference in the meaning between the adjective and the adverb that is derived with -*ly*.

(11) a. hard-*ly*

b. cold-*ly*

Specifically in the examples above the adverb *hardly* has a slightly different meaning than the adjective *hard*, while *coldly* can only be used in a metaphorical sense according to Plag. These semantic changes speak against the theory that -*ly* should be classified as an inflectional suffix. The literature on -*ly* is vast and a lot of time can be spent analysing this topic. In Hay's (2002) experiment on the relative frequency effect the data contained complex words with the -*ly* suffix and every affix that was used in the experiment was considered derivational. Since the experiments that were conducted for this thesis are similar to hers, the affix will be classified as derivational, albeit with a certain degree of uncertainty.

Summarising derivation in English, this field has considerably more variety than inflection and it is considered to be as rich as other European languages like French and German

(Carstairs-McArthy, 2002). This has a clear effect on the 'transparency' of the language, since semantic opacity is a property often found in derivation, but rarely found in inflection.

2.2.2 Word-Formation Processes

The complex words that were formed in the previous section are a combination of two elements, a root and an affix; prefix, suffix or infix. In English an essential feature of affixes is their combining power not only with a root but with other affixes as well. This results in a complex word that can have more than two morphemes. The word *helplessness* is composed of the root *help* and the suffixes -*less* and -*ness*. Multiple affixes can appear in a complex word. The word-formation process in which one or more affixes are attached to a root/stem is called affixation and it is the most common word-formation process in English.

Word-formation is not limited to affixation in English, since the language has a variety of processes to form words. One of these processes is back formation or reversion, which can be considered the exact opposite of affixation. Instead of adding an affix to a root, a shorter word is derived from a longer word by removing an imagined affix (Plag, 2003 & Crystal 2003). A characteristic example is the word *edit*, which is derived from the word *editor* by removing the suffix -*or*.

Another word-formation process, which is very common in English, is conversion, the process of making a new word from an existing word by changing the category of a part of speech, with the morphemic shape of the original word remaining unchanged, i.e. without the addition of an affix. The noun *bottle* is converted to the verb *(to) bottle*, while the adjective *empty* is converted to the verb *(to) empty*.

Compounding or composition is another major word-formation process in English, the most productive according to Plag (2003) and the most controversial in terms of linguistic analysis. A compound is a unit consisted of more than one lexical stem. The lexemes that compose the word act as a single item, which has a single main stress and it is used grammatically as a unit. For example the word *earthquake* has the stems *earth* and *quake* as its constituents and is used as a different lexeme than its composing parts. In compounds, and in general every complex word, one of the constituents will act as the

head of the word, i.e. the most important unit in the structure and it will define the word class of the compound. If the head will be a noun, then the compound will be a noun as well, e.g. *run* (V) + *way* (N) form the compound *runway* (N). In English the head is almost always on the right-hand side of the compound, which led Williams (1981) to name this phenomenon the right-hand head rule. There has been some criticism on this rule, since some examples show that the head does not appear to be on the right of the structure, e.g. *en-* (V) + *able* (Adj) form the word *enable* (V).

Contraction or shortening is the process, with which a word is formed by removing parts of a word. This can be achieved by clipping syllables, e.g. *ad* from the word *advertisement*, or by clipping the initial letters of a word group to form an acronym, e.g. *USA* from *United States of America*. Acronyms that are based on orthography are called blends (Plag, 2003). With blending new words are formed by combining two or more sounds or morphemes of several words. An example of a word formed by blending is *smog*, which formed as an amalgamation of two different words *smoke* + *fog*. Plag (2003) argues that blending should not be classified as a special form of compounding, but rather as a case of non-affixational derivation. He explains that they are constrained by semantic, syntactic and prosodic restrictions, while they behave semantically and syntactically like copulative compounds, i.e. compounds with two semantic heads like *bittersweet*.

Additional types of word-formation are onomatopoeia, the process of forming a word by imitating or resembling sounds, e.g. *tick-tock* from the sound of a clock, and reduplication, the formation of words by doubling the stem of an already existing word or by having very similar structures (Crystal, 2003). The latter process is also called partial reduplication (Plag, 2003). Examples of full reduplication are rare, while partial reduplication is more common with the change of a single consonant or vowel between the first and the second constituent, e.g. *walkie-talkie*.

CHAPTER 3: Languages Under Examination – Modern Greek

3.1 History of Greek

Just like English, Modern Greek (henceforth Greek), also called Neo-Hellenic, is part of the Indo-European family of languages, albeit being in a different subdivision and an independent branch, the Hellenic languages. There exist several confining theories about the origins of the Greek language (Adrados, 2005), but most of them agree that the Indo-European language is its ancestor. Its first speakers were dated sometime near 3,000 B.C. in the Neolithic or Bronze era and are believed to have existed in the Greek Peninsula.

The first script for writing Greek was in the form of the Linear B syllabary used for the archaic Mycenaean dialect in the 16th century BC, which was recently deciphered in the 20th century. This makes Greek the language with the longest documented history of any language belonging in the same family. After the fall of that civilization there is not a recorded script of the language for about five centuries that has survived to this day.

The next development period for the language is the Ancient Greek era in the archaic and classical periods of the Ancient Greek civilization. The language was composed of many dialects, the most important of which were the Attic, Ionian and Doric dialects, each of them corresponding to the three main tribes of the Greek. The Attic dialect is especially notable for including the classical texts of Plato and Aristotle, while the Ionian dialect is said to be the continuation of the dialect Homer used for his epic poems.

Since the 4th century BC the Greek language had passed into the Koine Greek period. The language at that time was a fusion of the Attic and Ionian dialects and became the first common Greek dialect, as well as the lingua franca of the time across the Mediterranean Sea and the Near East. Alexander the Great and his successors spread the language to many regions from Egypt to the depths of India during the Hellenistic colonisation. With the establishment of the Roman Empire, Koine Greek heavily influenced the Latin language, forming an unofficial bilingualism across the empire.

The continuation of Koine Greek was the Medieval Greek, also known as Byzantine Greek. It is recorded from the beginning of the Middle Ages, around 330 AC with the capital of the Roman Empire moving from Rome to Constantinople, up to the destruction

of the city and the demise of the Byzantine Empire in 1453. Up to this day this era serves as the connection point between the ancient and modern forms of the language, since the scripts at this period were heavily influenced by the Ancient Greek era, while the spoken language included many linguistic features that are present in the Modern Greek era.

The final recorded period of the language is the Modern Greek era, dated from the destruction of the Byzantine Empire in 1453 and it is still used in the modern age, albeit not completely identical to its form in the 15th century, since Greek is a language that keeps evolving. At the beginning of this period the Katharevusa form was considered the official language. It was a 'purified language' form according to Herzfeld (1987) and it was an attempt to denote the deep relation of the language with the Ancient Greek period without external influences. After 1976 the Katharevusa form was replaced with the Demotic dialect, which remains today's standard Modern Greek language, but certain words from Katharevusa still exist in the Demotic dialect.

Greek words are still entering languages like English as loanwords, especially in scientific fields. Along with Latin, Greek roots are still used in the formation of new words in many languages (Adrados, 2005) and they are the predominant sources of the international scientific vocabulary. Today it is officially spoken only in Greece and Cyprus and also spoken by minorities and immigrant communities in many countries across the world, including Germany, Canada and the United States of America.

3.2 Greek Morphology

Greek has a few similarities with English and many differences; The language is considered to be moderately rich in morphology (Ralli & Stavrou 1997, Melissaropoulou 2015) and a largely synthetic, inflectional language. Even though the complexity of the inflectional system has been considerably reduced compared to the language's Ancient Greek period, the language retains a degree of continuity in its morphological system.

Inflection and derivation in Greek work in a similar way to English, since both of them are synthetic and more specifically fusional languages. These languages use a single morpheme to denote multiple grammatical, syntactic and semantic features. From the many sub-categories of affixes analyzed in the previous chapters, Greek includes only prefixes and suffixes as well. Prefixes are limited in Greek and only four are recorded by

Ralli (2005), three of which have their roots in Ancient Greek (*a-*, *dys-* & *ef-*) and the remaining prefix *kse-* was formed during the Medieval Greek period. Just like English, Greek prefixes are only derivational and not inflectional contrary to Ancient Greek. In Ancient Greek the inflectional prefix *e-* was used for the formation of the past tense in verbs. Today its use is strictly related to the word's stress and does not have any inflectional purposes, since it is not mandatory in the formation of the past tense (Kaisse, 1982).

(12) a. έ-παι-ξ-α

'e *paik* *s* *a*

(stress element) play (aspect) ed (first person singular)

'I played'

b. παίξαμε

'paik *s* *ame*

play (aspect) ed (first person plural)

'we played'

In example 12a the first person singular form of the verb *paizo* (=to play) in the past tense requires the *e-* as a syllable augment in contrast with example 12b. The verb's first person plural form in the past tense already has three syllables and does not require the *e-* to form the word. The topic of the stress element *e-* has been discussed extensively in the literature and some views like Drachman & Malikouti-Drachman (1994) consider *e-* a circumfix, a discontinuous affix, whose phonetical components are not sequential within a structure. Ralli (2005), on the other hand, argues that in Ancient Greek, tense was indicated by the prefix, but in Modern Greek the inflectional suffixes are responsible for indicating it and since the appearance of e- is not systematic, the *e-* should be considered a stress element.

As far as the suffixes are concerned they can be either derivational or inflectional, as in the case of English, and they are more numerous than the prefixes in both languages. Especially in Greek the amount of suffixes that are recorded is greater than English.

Additionally, Greek shares the majority of its word-formation processes with English. One of the most common processes is affixation, which is the most common case of word formation in this language as well. Multiple affixes can exist within a complex word, but there are some constraints as to which affixes can be attached to certain stems. For

example the derivational suffix *-ar-* is attached to borrowed words in order to form a verb, e.g. *park-ar-o*, the act of parking.

Compounding is another major word-formation process in Greek and it is extremely productive in the language. The basic notion of this process is that same as English, two separate words are combined to form a new unit with every characteristic that a word has, such as the single stress that makes the compound a phonological word. Compounds in Greek do not always get their meaning from their components. If we examine the compound *ela'fromjalos*, its components are *ela'fros* (=light) and *mjalo* (=mind), yet the word's meaning is not 'he who has a light mind' but it holds the meaning of *careless*, which is derived metaphorically from its components.

Ablaut is another word-formation process that forms a derivative by changing the vowel in the word's stem, but it is not quite productive in the language, since its usage is significantly restricted around certain stems who originate from Ancient Greek, as seen in the example below, where the vowel in the verb's stem *'tr̲eho* (=run) changes in the noun's stem *tr̲o'hos* (=wheel)

(12) a. τρέχ-ω
treh *o*
run (1st person singular)
'I run'

b. τρ*o*χ-ός
troh *os*
wheel (nominative, singular)
'wheel'

At this point, however, it is necessary to explain the many differences between the two languages and some parameters that will affect the experiments. Even though, both languages share the same categories of affixes, derivational prefixes and inflectional and derivational suffixes, only the derivational affixes will concern us for this project, as the inflectional suffixes will be excluded from the experiments.

The reason that the inflectional affixes will not be of interest in this thesis is because of the richer morphology that Greek has, when compared to English, a language with a rather simple morphology as it was explained in the previous chapter. In certain languages such

as Greek, a word's stem is a bound morpheme, meaning that it has to be combined with an inflectional suffix to form a new word. These suffixes provide information about case, gender, number and tense. In languages like English, however, a stem can also be a word by itself, due to the language's small amount of inflectional morphemes.

(13) a. καρέκλ-α

ka'rekl *α*

chair 0 (nominative/accusative-singular)

'chair'

b. καρέκλ-ες

ka'rekl *es*

chair s (nominative/accusative)

'chairs'

c. καρεκλ-ών

kare'kl *on*

chair of the (genitive-plural)

'of the chairs'

(14) a. παίζ-ω

'pez *o*

play I

'I play'

b. παίζ-ουμε

'pez *ume*

play we

'we play'

c. θα παίξ-ω

tha *'pek* *s* *o*

will play (aspect) I

'I will play'

In the examples above, two stems are seen along with an inflection suffix attached to them. The stems *karekl-* and *pez-* are bound morphemes, meaning that they cannot appear in speech without the addition of a suffix. Examples 13a, 13b and 13c have the inflection suffixes *-a*, *-es* and *-on* respectively attached to the stem *karekl-*, which means *chair*. Adding the *-a* suffix creates the noun *ka'rekla* (=chair). In addition, the suffix carries

information about the gender of the noun (*ka'rekla* is a feminine gender), the number (*ka'rekles* means chairs) and the case (nominative or accusative case, depending on the article that accompanies the nouns, in 13a and 13b. Genitive case in 13c). The inflectional suffixes attached to nouns in Greek always carry information about the gender, the number and the case. Adjectives share this property with nouns.

In the case of verbs as seen in examples 14a, 14b and 14c the suffixes give information about the number, tense and aspect. The *-o* in *pezo* (14a) signals that the verb is in first person singular in the present tense. Unlike in English, a subject pronoun is not obligatory, since the information is provided by the suffix. It must also be noted that in the case of example 14c *tha 'pekso* (=I will play) we have a case of future tense, which is expressed by a periphrastic construction. The *k* is met in *pek-*, which is an allomorph to *pez-,* while the *s* is responsible for the aspect. Every person (first, second, third) number (singular, plural) tense (past and non-past) and aspect (perfective and imperfective) in Greek has its information given by a suffix or a combination of such (14c).

The aspects of Greek morphology considered above regarding the nouns, adjectives and verbs would make the relative frequency phenomenon substantially more difficult to examine compared to English. If English has two different forms in the lexeme RUN for the present tense (*run* & *runs*), Greek would have six different forms (*treh-o*, *treh-is*, *treh-i*, *treh-ume*, *treh-ete*, *treh-un*) one for each person in both numbers. For the past tense English would have one form (*ran*), Greek would again have six forms. When trying to find the total amount of instances a complex word is met in a corpus for English (e.g. *enrage*), it will not be as consistent as its Greek counterpart. *Enrage* may as well be in first or second person in singular or plural, or in third person plural, however for the Greek word we would need to search five different forms in its respective corpus. In conclusion, examining this phenomenon in Greek demands caution, since overlooking even a single form of the lexeme could bring imbalance to its frequency values.

A further complication is that Greek derivatives often involve more than one derivational processes. Derivative words in Greek are consisted of the affix along with the stem of the word. It is not rare, though, for the stem to additionally undergo through the process of ablaut, the process of forming a derivative by changing the vowel in the word's stem. This would mean that when trying to break down a complex word, a Greek native speaker may

have to remove the derivational prefix or the suffix and then perform an ablaut to the stem of the word, whereas an English speaker would only need to remove the affix.

(14) a. μασκ-ο-φόρ-ος
mask o' for os
mask (synthesis index) bring (nominative-singular)
'masked'
b. mask-ed

In examples 14 and 14b the same word in Greek and English is given. While in English it is relatively easy to decompose the word *masked* to its components, *mask* and *-ed*, we cannot say that it is just as convenient to do the same for 14a. On top of the word being a compound, the second component of the word *-foros* is derived from the verb *fer-o* (=to bring). The stem has gone through ablaut and was turned from *fer-* to *for-* before adding the inflection suffix *-os* and then be combined with the first component to form the word *masko'foros*, i.e. he who brings (has) a mask.

The Greek native speakers are able to distinguish the meaning of these words, however in order for the experiments to be as similar as possible across the two languages, the examples chosen for Experiment 2 will have to be easily disassembled to prefix/suffix + stem without any other process intervening.

CHAPTER 4: Methodology

4.1 Word Pairs

The set-up of the experiments was providing the participants with word pairs and asking them which of the two words is more complex according to them. Thirty (30) pairs of words were composed for each experiment, for a total of sixty (60) word pairs across the two experiments, of which twenty (20) were prefixed pairs and another twenty (20) suffixed pairs. The remaining twenty (20) word pairs included in the study were filler.

The word pairs under examination were chosen based on their frequency, where one of them had a considerably higher relative frequency than its base and the other had a considerably lower relative frequency than its base. The frequency information for the English words was obtained by the Corpus of Contemporary American English (C.O.C.A.) database, which includes more than 520 million words, while for Greek words the Corpus of Modern Greek (C.O.M.G.) database was chosen with an approximately 35,7 million words in its files. The uneven amount of words that each database has, is explained due to the unfortunate lack of multiple online databases of Greek corpora when compared to English.

The members of the word pairs share the same affix and were matched carefully in order to depict a resemblance in stress pattern, syllable count and junctural phonotactics whenever this was possible. In the case of Greek the gender and the inflectional suffixes factors were additionally taken into consideration,. The filler pairs included affixed and pseudo-affixed words.

In the Tables below the words on the left have a higher frequency than their base, as opposed to the words on the right, which show a lower frequency than their base. Since the relative frequency is lower for the words on the right, it is therefore expected that these words will be considered more complex by the participants. Tables 1 and 3 list the English and Greek prefixed word pairs respectively and Tables 2 and 4 for the suffixed ones. The filler pairs are shown in Table 5 for both English on the left side and Greek on the right side.

The Greek words are followed by transcription to the Latin alphabet and translation to English. Furthermore, the frequency values for the Greek suffixed stimuli need to be

specially mentioned. The derivational suffixes are followed by inflectional ones, which provide us with important information as mentioned above. For that reason, it was required to search the word examples in the database for every gender, number, person and case possible for both the derivatives' and the bases' frequency. The amount shown in the tables is the accumulated amount, however only the nominative case of the masculine gender in singular number is shown in the table for nouns and adjectives, while for the verbs the first person in singular number in present tense.

The order of the word pairs along with the words within these pairs were randomized during the experiments.

Table 1 English Prefixed Stimuli

Word A	Frequency	Base Frequency	Word B	Frequency	Base Frequency
engage	16,622	785	enrage	127	8,813
inanimate	712	655	inadequate	6,334	11,835
irreplaceable	568	289	irremovable	8	1,023
irreversible	1,189	744	irredeemable	68	112
refurbish	268	11	rekindle	344	823
revamp	608	283	recycle	1,515	16,182
unbelievable	5,238	3,142	unacceptable	4,000	11,307
uncanny	1,602	467	unmanly	84	1,092
unforgivable	423	140	unfavorable	1,363	6,545
untouchable	576	101	unteachable	23	316

Table 2 English Suffixed Stimuli

Word A	Frequency	Base Frequency	Word B	Frequency	Base Frequency
abasement	32	21	statement	37,826	331,685
abysmal	555	4	facial	5,512	183,490
exactly	72,130	12,288	chilly	2,700	8,193
emotional	30,284	8,967	logical	7,119	10,631
finally	106,115	66,069	generally	41,680	135,747
frequently	26,174	12,832	meekly	396	834
hapless	978	594	ruthless	2172	9,916
national	216,431	80,247	musical	24,838	142,178
slimy	856	867	dirty	15,037	16,153
virility	246	5	futility	1,033	2,055

Table 3 Greek Prefixed Stimuli

Word A	Frequency	Base Frequency	Word B	Frequency	Base Frequency
ανάδειξη a'naδiksi emergence	1,338	71	ανακατάταξη anaka'tataksi rearrangement	14	512
ανεπαρκής anepar'kis insufficient	161	195	ανέτοιμος a'netimos unprepared	87	1,650
απογοήτευση apogo'itefsi disappointment	676	278	αποκάλυψη apo'kalipsi discovery	867	1,826
διασυρμός δiarsi'mos calumniation	42	37	διάδρομος δi'aδromos walkway	68	2,156
δυσοσμία δisos'mia stench	57	105	δυσπιστία δispis'tia disbelief	194	1,446
εκτέλεση ek'telesi execution	1,632	178	εκδήλωση ek'δilosi demonstration	3,669	4,517
εξασφάλιση eksa'sfalisi security	629	489	εξάντληση e'ksadlisi depletion	137	219
προσγείωση pros'jiosi landing	238	4	πρόσκρουση 'proskrusi collision	62	100
συμμόρφωση si'morfosi compliance	210	201	συμπίεση si'biesi compression	67	2,544
υπογραφή ipoγra'fi signature	2,040	452	υποθήκη ipo'thiki mortgage	57	83

Table 4 Greek Suffixed Stimuli

Word A	Frequency	Base Frequency	Word B	Frequency	Base Frequency
αντάλλαγμα a'dalaγma exchange	649	219	βρόντηγμα 'vrodiγma rumbling	0	39
δριμύτητα δri'mitita severity	1,552	23	εγγύτητα e'gitita imminence	49	296
εγκλωβισμός egklovi'smos lock-in	17	10	υποβιβασμός ipoviva'smos demotion	71	98
ηλιακός ilia'kos solar	836	875	σταδιακός staδia'kos gradual	2,345	2,923
κομβικός komvi'kos nodal	192	170	ζωικός zoi'kos animalistic	4	292
κομμωτής komo'tis hairdresser	57	5	διαδηλωτής δiaδilo'tis protester	1,826	2,156
οικειότητα ici'otita intimacy	57	37	τελειότητα teli'otita perfection	75	429
χάραμα 'harama dawn	34	4	παίδεμα 'peδema bedevilment	1	9
χλοερός hloe'ros grassy	48	44	φοβερός fove'ros fearsome	352	994
ψητός psi'tos roast	24	6	στητός sti'tos upright	1	6

Table 5 English & Greek Filler Stimuli

ability	invisibility	αναλογικός analoγi'kos analogical	ανάμεικτος a'namiktos mixed
adjective	effective	άνετος 'anetos comfortable	ανήσυχος a'nisixos alarmed
default	debug	διάβολος δi'avolos devil	διάλογος δi'alογos dialogue
defend	deblend	έκτος 'ektos sixth	έκταση 'ektasi acreage
engine	enjoin	εξωτερικός eksoteri'kos external	εξοργιστικός eksorγisti'kos infuriating
family	heavily	κόκκινος 'kokinos red	πέτρινος 'petrinos stony
mission	misfunction	κρίνος 'krinos lily	θεατρίνος 'thea'trinos luvvy
nevertheless	pointless	πανικός pani'kos panic	μηχανικός michani'kos mechanic
pedestrian	librarian	ποντικός ponti'kos mouse	βασιλικός vasili'kos royal
quality	vitality	σεισμός si'smos earthquake	συλλαβισμός silavi'smos spelling

4.2 Participants

Twenty (20) adult native speakers were selected for each of the two experiments, forty (40) participants in total, all of which voluntarily participated after they were asked to. For the first experiment all of the participants were native English speakers, while for the second experiment they were native Greek speakers. The participants gave their consent in verbal form to have their results used for this thesis, after being reassured that their anonymity would be kept and only their initials would be used for reference.

An effort was made to have the same number of male and female participants (although the participants' gender was not expected to influence the results). For Experiment 1 in English there were nine (9) male and eleven (11) female participants and for Experiment 2 in Greek twelve male (12) and 8 female (8) participants, for a total of twenty-one (21) males and nineteen (19) females.

All of the participants shared a similar educational background with all of them being university students, albeit in different fields of science. However, none of them was associated to linguistics and had little to no knowledge of the field. The participants' age varied between twenty (20) and thirty (30) years old.

4.3 Materials

The first experiment (on English) was carried out using the online survey platform SurveyMonkey.com.

Due to the fact that none of the online survey platforms supports the Greek alphabet and since proficiency in English was not a requirement for the Greek participants, the second experiment had to take the form of a questionnaire, in order for the Greek subjects to be able to participate in the experiments.

The possible differences in the experiments were taken into consideration as the Greek participants would have the convenience of re-examining their answers at any point during the filling of the questionnaire. For this reason, the English participants were told that they had the option to re-examine their answers before submitting them.

CHAPTER 5: Experiment 1 - English

5.1 Procedure

After meeting with the participants, either via an online video call or in person, and verbal consent was provided, the following instructions were given to them in written form as part of the experiment. The instructions were based on those in Hay's (2000) experiment, albeit with different examples and slightly modifying the text in an attempt to be more accurate for the participants.

The subjects completed the experiment at their own pace after been asked if they had any questions prior writing down their answers.

An example of the experiment along with the instructions can be seen in the Appendices section.

5.2 Results

Any subjects that did not consistently distinguish the pseudo-affixed and affixed filler word pairs were not included in the analysis. Specifically, one participant was excluded, since his/her results were heavily inconsistent, as shown by the small amount of correct answers in the filler questions. Another participant systematically gave the exact opposite expected answer, which shows a misunderstanding in the instructions on their behalf. The same issue was spotted in Hay's original experiment, who named it "a rather terminological confusion than a conceptual one". Their data was included in the analysis, after the answers were reversed.

Therefore nineteen (19) participants were analyzed. The data show that the forms whose frequency is lower than their bases were consistently rated as more complex than the forms whose frequency is higher than their bases. This is clearly shown in Table 6, where roughly only 10.78% of the total answers were in favour of the word with a higher frequency than its base compared to the 89.21% of the total answers that rated the words with a lower frequency than its base more complex.

The above statement is true for both prefixed (11.57% to 88.42% in favour of the word with lower-than-the-base frequency) and suffixed pairs (8.94% to 91.05% for the word with lower-than-the-base frequency), roughly the same tendency for all affixed pairs with the suffixed ones having a slightly higher percentage.

Table 6: Experiment 1 Results

Word A	Answer Rate	Word B	Answer Rate
engage	5.26%	enrage	94.73%
inanimate	15.78%	inadequate	84.21%
irreplaceable	15.78%	irremovable	84.21%
irreversible	10.52%	irredeemable	89.47%
refurbish	10.52%	rekindle	89.47%
revamp	0.00%	recycle	100.00%
unbelievable	21.05%	unacceptable	78.94%
uncanny	10.52%	unmanly	89.47%
unforgivable	21.05%	unfavorable	78.94%
untouchable	5.26%	unteachable	94.73%
abasement	10.52%	statement	89.47%
abysmal	5.26%	facial	94.73%
exactly	15.78%	chilly	84.21%
emotional	5.26%	logical	94.73%
finally	0.00%	generally	100.00%
frequently	15.78%	meekly	84.21%
hapless	10.52%	ruthless	89.47%
national	0.00%	musical	100.00%
slimy	15.78%	dirty	84.21%
virility	10.52%	futility	89.47%

CHAPTER 6: Experiment 2 – Modern Greek

6.1 Procedure

After meeting with the participants and their verbal consent was provided, the following instructions were given to them in written form as part of the experiment. An important addition to this experiment's instructions was clarifying to the subjects the meaning of the term 'complex'. The word *complex* has more than one meaning in Greek and the participants could be confused as to what was being asked from them, so they would give as accurate answers as possible. The instructions were given in Greek and were based on the instructions of the previous experiment, however for convenience of the readers the instructions were directly translated in English. For a detailed example of the experiment in Greek, refer to the Appendices section at the end of the thesis.

The subjects completed the experiment at their own pace after been asked if they had any questions prior writing down their answers.

6.2 Results

As with the first experiment any subjects that did not consistently distinguish the pseudo-affixed and affixed filler word pairs were not included in the analysis, however all participants systematically answered correctly in at least 80% of the filler word pairs. Similarly with a case of the first experiment conducted in English, a participant consistently provided the exact opposite expected answer even in filler pair, which shows a misunderstanding in the instructions on their behalf.

Therefore all twenty (20) participants' answers were analyzed. The data shows once more that the forms whose frequency is lower than their bases were mostly rated as more complex than the forms whose frequency is higher than their bases. This is clearly shown in Table 7, where only 30.25% of the total answers were in favour of the word with a higher frequency than its base compared to the 69.75% of the total answers that rated the words with a lower frequency than its base more complex.

The above statement is true for both prefixed (35.5% to 64.5% in favour of the word with lower-than-the-base frequency) and suffixed pairs (25% to 75% for the word with lower-than-the-base frequency), about the same tendency for all affixed pairs with the suffixed ones having a considerably higher percentage.

Table 7: Experiment 2 Results

Word A	Answer Rate	Word B	Answer Rate
a'naδiksi	10,00%	anaka'tataksi	90,00%
anepar'kis	65,00%	a'netimos	35,00%
apogo'itefsi	50,00%	apo'kalipsi	50,00%
δiasir'mos	35,00%	δi'aδromos	65,00%
δisos'mia	25,00%	δispis'tia	75,00%
ek'telesi	20,00%	ek'δilosi	80,00%
eksa'sfalisi	50,00%	e'ksadlisi	50,00%
pros'jiosi	45,00%	'proskrusi	55,00%
si'morfosi	20,00%	si'biesi	80,00%
ipoγra'fi	35,00%	ipo'thiki	65,00%
a'dalaγma	10,00%	'vrodiγma	90,00%
δri'mitita	30,00%	e'gitita	70,00%
egklovi'smos	40,00%	ipoviva'smos	60,00%
ilia'kos	40,00%	staδia'kos	60,00%
komvi'kos	10,00%	zoi'kos	90,00%
komo'tis	5,00%	δiaδilo'tis	95,00%
ici'otita	45,00%	teli'otita	55,00%
'harama	25,00%	'peδema	75,00%
hloe'ros	20,00%	fove'ros	80,00%
psi'tos	25,00%	sti'tos	75,00%

CHAPTER 7: Discussion

7.1 Analyzing the Results

The first question that must be addressed in this chapter is whether the results of the two experiments support the hypothesis that there are certain factors which affect the decomposability of affixed words is confirmed or not. The answer is that they do as the majority of the participants' answers inclined towards the word that had a lower frequency than its base and thus they rated them more complex compared to the other word. In the first experiment, on English, about 89% of the total answers favoured the affixed word that had a lower frequency than its base. In the second experiment, on Greek, the number ranges to roughly 70%.

The results therefore confirm the existence of the relative frequency effect as argued for by Hay (2000). Indeed, the results of the experiment on English show this effect even more strongly than in Hay's original experiment.

Second, and more importantly, the results reveal the subjects' tendency in the rating of complex words. As expected, the participants in their majority chose the affixed word with the lower relative frequency, as found in the corpus database, as being the more complex of the pair. This means that these words are more easily decomposed according to the subjects than the affixed words having a higher relative frequency.

The results of the experiments suggest not only that the relative frequency effect exists, but also that it is one of the factors that affects the decomposability of affixed words down to their components. This statement is true for both English and Greek, although it must be mentioned that for the second experiment the results have an important gap of about 20% compared to the first experiment's results. This will be discussed later in detail when the cross-linguistic factor is examined.

Another factor that was hypothesized to affect the decomposition of complex words was the affix' position in relation to the stem, i.e. whether the affix in the complex word would be a prefix or a suffix. Seemingly the difference in the answer rates between word pairs that contained either of the two affixes is relatively low. In the first experiment 88.42% of the total answers in the prefixed pairs favoured the low-frequency word, as opposed to the

91.05% in the suffixed pairs. Following the second experiment, 64.5% of the answers were noted for the low-frequency word in the prefixed pair and 75% for the suffixed pair. We observe another potential factor in the decomposability of affixed words. However, the majority of both affixed pair categories is in agreement with the relative frequency effect even though not on the same rate. In English the difference is miniscule and in Greek we have a 10% gap.

In both languages, though, we observe that the suffixes have a higher percentage in the answers, showing a tendency to be more frequent than the affixed word's base. So how do the suffixed forms differ from the prefixed forms? Hay justified this difference in regards to the onsets of the base and its derived form. The onset of the base and its suffixed form occurs simultaneously, however the prefixed form's onset is temporally prior to the onset of the base. This is speculated by Hay to affect the decomposability of affixed words, which is naturally connected with the relative frequency effect. The difference is relatively small, however, by which we can conclude that although the suffixed forms are more easily decomposed than the prefixed forms, both can be decomposed by speakers in a considerably high frequency.

Finally, we must not forget about the cross-linguistic factor. The experiments proved that the relative frequency effect and the position of the affix in the word are factors that affect the decomposability of complex words in both languages. However, there is a mismatch in the data between English and Greek.

The lower frequency words for the first experiment were rated complex in almost 90% of the total answers, but for Greek roughly 70%. While there is only a 3% difference in the recorded answers for the prefixed and suffixed pairs in English, we notice a 10% gap for the pairs in the second experiment, a significant gap between the two languages. Why does Greek show such a huge difference compared to English?

7.2 The Cross-Linguistic Factor

The answer to the question as to why Greek has different percentages compared to English in the above experiments possibly lies in the very morphology of the languages and the complexity of the affixed words. As was already stated in §2: Languages Under Examination, ***E***nglish has relatively little morphology as compared to Greek. The complexity of a Greek word's composition with the inflectional suffixes and the abundance of bound stems is probably one of the reasons for the occurrence of this phenomenon. We must not eliminate, however, the possible existence of more factors that affect decomposability, which vary between the two languages, and to discover them we must carefully examine the words used in the experiments.

Despite the discovery and the confirmation of certain factors affecting decomposability, one could logically wonder; why are these factors not absolute? Even in English where a crushing 90% of the total answers inclined towards the low-frequency word in rating it more complex, there remains another 10% that had a different opinion. In Greek this number is even higher, as explained in the previous section. Why is there not an almost conclusing percentage in the answers? Why did some participants choose the other word of the two?

One possible explanation could be that the participants were uncertain of which word appeared to be more complex and they randomly chose between the two words. However, that cannot be justified by the collected data. Both current and Hay's original experiment agree in the participants' tendency to choose the low-frequency word. For example, in the word pair *engage – enrage*, only one participant chose the higher-than-the-base frequency word *engage* as being more complex. If both words were equally complex, then we would expect a somewhat equal distribution of answers, something that did not occur. *Enrage* was chosen by 18 participants as the more complex word of the pair.

When the difference is considerably small as shown in the first experiment's results, we can narrow down the explanation to each subject's personal perception of how complex a word really is. But as a general rule, relative frequency and the position of the affix in the complex word give us a hint to which word can be more complex.

In the case of Greek I ought to give a more convincing explanation about the difference in the data. In the English experiment three out of the twenty word pairs had an 100% preference in the low frequency words, while the remaining pairs heavily favoured the low frequency affixed form. For the second experiment a considerable amount of word pairs showed mixed preferences between the low and high frequency words and one word pair *anepar'kis* (=insufficient) – *an'etimos* (=unprepared) had the completely opposite expected result. The reason for this outcome is in the frequency itself. Both words of the pair above appeared less times than their bases in the corpus database. The first word *anepar'kis* has an almost even number of instances in the database, whereas the second one *an'etimos* almost diminishingly compared to its base. As a result, it was expected that the word that appeared the most in the database would also be rated less complex. The reason that this example was chosen was because of my attempt to depict a resemblance in stress pattern, syllable count and junctural phonotactics whenever this was possible and for the lack of a better candidate, the word *anepar'kis* was chosen. Moreover, similar cases like the aforementioned were included in Hay's original experiment for English in order to find the most suitable pair word possible. In the word pair *agility – fragility* in Hay's experiment, both words have lower frequency values than their base with the sole difference that *fragility* has far less instances in the database than its base compared to *agility*. The results were favourable in her case but unfortunately not in this project.

One factor that could play an important role in the puzzle of cross-linguistic differences is the stress for the case of Greek. Unlike English, where almost every prefixed and suffixed forms hold the stress in the same syllable as the base form, in Greek a property of most affixes replaces the stress to a previous or upcoming syllable in the word. An evident example would be the word pair *pros'jiosi* (=landing) - *'proskrusi* (=collision). Even though the second word '*proskrusi* has a lower frequency than its base '*krusi* participants seemed to also prefer the other option in the pair. Contrary to the base *'krusi* where the stress is moved to the previous syllable because of the *pros-* prefix, the word *pros'jiosi* holds the stress at the very same place as the base *'jiosi*. This potentially confuses the speakers, as they will head for the word that is immediately accessed in the lexicon.

One final possibility to the answer's inconsistency is that the subjects failed to perceive the connection of the affixed form and base due to the slight semantic differences of the form. Specifically, by examining the word pair *eksa'sfalisi* (=security) and *e'ksadlisi* (=depletion) we observe that the participants were divided in half for this example, despite the original

prediction that *e'ksadlisi* would receive the most answers. With a closer inspection to the word itself, however, we will notice a slight difference between this form and its base. *'Adlisi* holds the meaning of sucking out a liquid and *e'ksadlisi* the depletion of something, not necessarily a liquid and can also be used in metaphors. *Eksa'sfalisi* on the other hand holds the meaning of security and *a'sfalisi* securing something via an insurance. The latter word holds a deeper connection to its base despite having a lower frequency, which goes against the theoretical background as it was explained in *Chapter 1: Decomposability and Relative Frequency*. Normally when the derived form of a base is more frequent than the base itself, then it would be more likely to undergo a semantic drift, something that does not happen in some of the word pairs in Greek.

In summary, even though the main hypothesis that certain factors affect the decomposability of affixed words was confirmed for the case of English, specifically the relative frequency effect and the affix' position, there seems to be additional factors for languages with different morphology than English, as in the case of Greek. Relative frequency and the affixes in the word are still in effect when it comes to decomposability and they seem to be the more dominant factors. Following these factors we have the stress movement in the affixed forms and the semantic drift that is sometimes irrelevant to relative frequency. This leads us to believe that there may be a hierarchy in the factors that are responsible for this phenomenon, starting with the relative frequency effect, then affix positioning and concluding with stress movement and semantic drift.

CHAPTER 8: Conclusion

To sum this thesis up, decomposability is a property of affixed words, i.e. words that are a combination of a stem and an affix. This property is about breaking down this complex form down to its constituent parts in order to clarify its meaning. Every affixed word is decomposable and every native speaker has the ability to break down these complex words. The goal of this thesis was to determine which factors are responsible for assisting this process. Additionally, an attempt was made to examine if these factors are applied in a cross-linguistic level. This would be tested by examining two different languages.

After stating the specifics of the English and Greek languages and describing their various differences, two experiments were conducted, one in each language, in order to confirm the hypothesis that relative frequency and the position of the affixes in a complex word can facilitate its decomposing down to its constituent parts and to examine any differences between the two languages as well as the existence of possible additional factors.

The experiments included word pairs in English and Greek respectively that were carefully chosen from online corpus databases for each language. In each pair one word had a higher frequency in the databases than its base and the other word a lower frequency. The word pairs included both prefixed and suffixed forms of bases and both words were as similar as possible in regards to stress pattern, syllable count and junctural phonotactics.

After the participants completed the experiments and the collected data was analyzed, the hypothesis of the relative frequency effect and the position of the affixes affecting the decomposability of affixed words was confirmed. However, the data from the second experiment in Greek revealed that in a language with a different morphology than English other factors may underlie the process. Specifically, the stress movement in an affixed word due to the property of some affixed in Greek along with a semantic drift not necessarily occurring to words that have a higher frequency than their base but also in low frequency derived words play an important role in decomposing complex words. According to the data, the majority of the word pairs in Greek were still influenced by the relative frequency effect and the position of affixes and in a smaller scale by semantic drift and stress movement.

Before this project is finished, I will make a few comments regarding the question that was set a few times during the thesis, about how this theory could be applied to other languages with different morphology than English and Greek. English is an analytic language with a characteristic low ratio of morphemes to words. Greek is a fusional language, where morphemes are not readily distinguishable from the root or among themselves. What would happen if we applied this theory to an agglutinative language such as Japanese, where every morpheme (stem and affix) remains unchanged after being composed in the word? How would such a language differ from English in decomposability? Another example would be a polysynthetic language like the Siberian language Ket, where words are extremely long and are composed of many morphemes.

Taking everything that was mentioned into consideration, we can understand how many similarities and differences English and Greek have both as languages and as far as decomposability is concerned. The relative frequency effect brings different results in languages with different morphology than English. On the context of studies about decomposability, the next step could be to determine if there are other factors that this thesis has overlooked and examine other languages belonging in families other than the Indo-European family. Further examination is required to find out each language's differentiations and to answer questions about the universality of the relative frequency effect.

References

- **Aas Barts** & **McMahon April**, 2008. *The Handbook of English Linguistics*. Language Arts & Disciplines. Wiley Publications.
- **Adrados Francisco Rodriguez**, 2005. *A History of the Greek Language: From Its Origins to the Present.* BRILL Publications.
- **Aronoff Mark** & **Fudeman Kirsten**, 2011. *What is Morphology?*. John Wiley & Sons Publications.
- **Bauer Laurie**, 2002. *English Word Formation.* Cambridge University Press.
- **Laurie Bauer**, **Rochelle Lieber** & **Ingo Plag**, 2013. *The Oxford Reference Guide to English Morphology.* Oxford: Oxford University Press
- **Carstairs-McCarthy Andrew**, 2002. *An Introduction to English Morphology.* Edinburgh: Edinburgh University Press.
- **Crystal David**, 2003. *The Cambridge Encyclopedia of the English Language.* 2nd ed. Cambridge University Press
- **Drachman Gaberell** & **Malikouti-Drachman**, 1994**.** *Stress and Greek Compounding*. In Phonologica 1992: 55-64
- **Hay Jennifer***,* 2001. *Lexical frequency in morphology: Is everything relative*? Linguistics. 39. 1041-1070. 10.1515/ling.2001.041.
- **Herzfeld Michael**, 1987. *Anthropology Through the Looking-Glass: Critical Ethnography in the Margins.* Cambridge University Press
- **Kaisse Ellen M**, 1982. *On the preservation of stress in Modern Greek*. Linguistics, 20, 59–82.
- **Marchand Hans,** 1971. *The categories and types of present-day English word-formation.* München: Verlag C. H. Beck. Second edition, 1969. Pp. x–xxvii, 1–545.. Journal of Linguistics.
- **Matthews Peter**, 1991. *Morphology*. 2nd edn. Cambridge: Cambridge University Press.
- **Melissaropoulou Dimitra**, 2015. *Modern Greek*. Edinburgh Handbook of Evaluative Morphology. Edinburgh University Press.
- **Pavol Štekauer** & **Lieber Rochelle**, 2005. *(The) Handbook of Word Formation.* Springer publications.
- **Plag Ingo**, 2002. *Word formation in English* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press, 2003. Pp. xiv+240. -. Journal of Linguistics. 41. 218 – 220

- **Ralli Aggeliki & Stavrou Melita**, 1997. *Morphology-Syntax Interface*. Yearbook of Morphology 1997. Ed G.E. Booij, Jaap van Marle. Springer Science & Business Media.
- **Ralli Aggeliki**, 2004. *Stem Based vs. Word Based Morphological Configurations.* Lingue e Linguaggio II, 241-275. Bologna: Il
- **Ralli Aggeliki**, 2005. *Morphology.* Patakis Publications.
- **Siegel, Dorothy C.**, 1974. *Topics in English morphology*. MIT Press.
- **Tesitelova Marie**, 1992. *Quantitative Linguistics*. Benjamins Publications.
- **Walter de Gruyter** 2000 *Morphology*. Ed. by Booij, Geert E. / Lehmann, Christian / Mugdan, Joachim In collab. with Kesselheim, Wolfgang / Skopeteas, Stavros
- **Williams Edwin** 1981. *On the notions "lexically related" and "head of a word"'*. Linguistic Inquiry 12: 245-274.

APPENDICES

APPENDIX 1

Experiment 1 Template

Complex Words

Instructions

The experiment you will participate in is about complex words.

A complex word is a word which can be broken down into smaller, meaningful blocks. In English, for example, the word "worker" can be broken down into two blocks: work and -er. The -er is a block which occurs at the end of many English words. In the example above, -er has been added to the word work to make a new, more complex word worker.

"Rework" is another example of a complex word in English. It can be broken down into re- and work.

Words, which are not complex, are called simple words. Here are some examples of simple words in English: Chair, brother, book, laugh. It is not possible to break down these words into smaller blocks and therefore are simple words.

In this experiment, you will be presented with pairs of complex word and you will be asked to decide which one do you think is MORE complex.
For example, "darkness" is very complex. It can be easily broken down into dark and -ness. "Business", however, is not quite so complex. While it is possible to break business down into busy and -ness it does not seem completely natural to do so. Business is complex, but not as complex as darkness.

Another example of a complex word is "disyoke". Even though you may never have heard the word disyoke before, you can understand its meaning, because it can be broken down into dis- and yoke. "Discard" is also complex – it can be broken down into dis- and card. But discard does not seem as complex as disyoke. We do not need to break discard into its parts in order to understand its meaning, and, in fact, it seems slightly unnatural to do so.

For each pair of words below, please read both words silently to yourself and then fill in the word you think that is more complex. It is very important that you provide an answer

for every pair, even if you are not certain of your answer. There is no right or wrong answer, just follow your intuition and provide your best guess.

Thank you very much in advance!

In the boxes below fill in the word you believe to be the more complex

1. facial – abysmal
2. inadequate – inanimate
3. mission – misfunction
4. librarian – pedestrian
5. frequently – meekly
6. generally - finally
7. family - heavily
8. dirty - slimy
9. unbelievable - unacceptable
10. ability - invisibility
11. uncanny - unmanly
12. deblend - defend
13. virility -futility
14. pointless - nevertheless
15. ruthless - hapless
16. unforgivable - unfavorable
17. logical - emotional
18. abasement - statement
19. recycle - revamp
20. untouchable - unteachable
21. quality – vitality

22. engage - enrage

23. default - debug

24. rekindle - refurbish

25. chilly - exactly

26. musical - national

27. irreversible - irredeemable

28. effective - adjective

29. engine - enjoin

30. irreplaceable - irremoveable

APPENDIX 2

Experiment 2 Template

Σύνθετες Λέξεις

Οδηγίες

Το πείραμα στο οποίο θα συμμετάσχετε σχετίζεται με τις σύνθετες λέξεις.

Σύνθετη λέξη είναι μια λέξη που μπορούμε να την σπάσουμε σε μικρότερα κομμάτια που είναι φορείς σημασίας. Στα Ελληνικά για παράδειγμα, η λέξη "πρασινωπός" μπορεί να χωριστεί σε δύο μικρά κομμάτια: "πράσινο" και "-ωπός". Το "-ωπός" είναι ένα κομμάτι που συναντάται στο τέλος αρκετών Ελληνικών λέξεων. Στο παραπάνω παράδειγμα, το "-ωπός" προστέθηκε στη λέξη "πράσινο" για να φτιάξουμε μια νέα, πιο σύνθετη λέξη "πρασινωπός".

Το "καταπράσινος" είναι ένα άλλο παράδειγμα σύνθετης λέξης στα Ελληνικά. Μπορούμε να το σπάσουμε σε "κατα-" και "πράσινος". Οι λέξεις, που δεν είναι σύνθετες, ονομάζονται απλές λέξεις. Μερικά παραδείγματα απλών λέξεων στα Ελληνικά είναι: χωρίς, σήμερα κτλ. Δεν είναι δυνατόν να σπάσουμε αυτές τις λέξεις σε μικρότερα κομμάτια, επομένως θεωρούνται απλές λέξεις.

Σε αυτό το πείραμα, θα σας δοθούν μερικά ζευγάρια λέξεων και θα σας ζητηθεί να αποφασίσετε ποια από τις δύο λέξεις είναι πιο σύνθετη.

Για παράδειγμα, η λέξη "εκατομμυριούχος" είναι πολύ σύνθετη. Μπορούμε να την σπάσουμε εύκολα σε "εκατομμύριο" και "-ουχος". Η λέξη "ευνούχος" όμως δεν είναι τόσο σύνθετη. Ενώ μπορούμε να σπάσουμε το "ευνούχος" στο αρχαιοελληνικό ουσιαστικό "ευνή" και το "-ουχος" δεν φαίνεται εντελώς φυσικό. Η λέξη "ευνούχος" είναι σύνθετη, αλλά όχι τόσο σύνθετη όσο το "εκατομμυριούχος".

Ένα άλλο παράδειγμα σύνθετης λέξης είναι το "υπόρραμμα". Παρόλο που μπορεί να μην έχετε ξανακούσει τη λέξη "υπόρραμμα", μπορείτε να καταλάβετε τη σημασία της, γιατί μπορούμε να τη σπάσουμε σε "υπο-" και "ράμμα". Το "υπόληψη" είναι επίσης σύνθετη λέξη – μπορούμε να τη χωρίσουμε σε "υπο-" και "λήψη". Δεν χρειάζεται να χωρίσουμε τη λέξη "υπόληψη" σε περαιτέρω κομμάτια για να καταλάβουμε τη σημασία της και μάλιστα φαίνεται λίγο περίεργο αν το κάνουμε.

Για κάθε ζευγάρι λέξεων που σας δίνεται παρακάτω, διαβάστε και τις δύο λέξεις σιγανά και τότε κυκλώστε τη λέξη που θεωρείτε ότι είναι πιο σύνθετη. Σας υπενθυμίζεται ότι με τον όρο "σύνθετο" δεν εννοούμε την πιο περίπλοκη ή την πιο δύσκολη λέξη ή την λέξη που πιθανόν να έχετε συναντήσει λιγότερες φορές, αλλά τη λέξη που μπορούμε να σπάσουμε στα κομμάτια που την αποτελούν πιο εύκολα από την λέξη που τη συνοδεύει.

Είναι πολύ σημαντικό να δώσετε μια απάντηση για κάθε ζευγάρι, ακόμα και αν δεν είστε σίγουροι για την απάντησή σας. Δεν υπάρχει σωστή ή λάθος απάντηση, απλώς ακολουθήστε το ένστικτό σας και δώστε την απάντηση που θεωρείτε εσείς σωστή.

Σας ευχαριστώ πολύ για το χρόνο και τη συμμετοχή σας!

Συμπληρώστε στα κουτιά τη λέξη που θεωρείτε ότι είναι πιο σύνθετη.

1. εξωτερικός - εξοργιστικός

2. διάλογος - διάβολος

3. συμμόρφωση - συμπίεση

4. διασυρμός - διάδρομος

5. αναλογικός - ανάμεικτος

6. διαδηλωτής - κομμωτής

7. υπογραφή - υποθήκη

8. ανάδειξη - ανακατάταξη

9. ανήσυχος - άνετος

10. ποντικός - βασιλικός

11. δυσπιστία - δυσοσμία

12. δριμύτητα - εγγύτητα

13. εκδήλωση - εκτέλεση

14. παίδεμα - χάραμα

15. ψητός - στητός

16. προσγείωση - πρόσκρουση

17. χλοερός - φοβερός

18. ανέτοιμος - ανεπαρκής

19. έκταση - έκτος

20. εξάντληση - εξασφάλιση

21. ηλιακός - σταδιακός

22. εγκλωβισμός - υποβιβασμός

23. κόκκινος - πέτρινος

24. αποκάλυψη - απογοήτευση

25. ζωικός - κομβικός

26. σεισμός - συλλαβισμός

27. οικειότητα - τελειότητα

28. κρίνος - θεατρίνος

29. βρόντηγμα - αντάλλαγμα

30. πανικός - μηχανικός

Experiment 2 Instructions Translated in English

The experiment you will participate in is about complex words.

A complex word is a word which can be broken down into smaller, meaningful **blocks**. In Greek, for example, the word *prasino'pos* (=greenish) can be broken down into two blocks: '*prasino (=green)* and *-o'pos*. The *-o'pos* is a block which occurs at the end of many Greek words. In the example above, *-o'pos* has been added to the word '*prasino* to make a new, more complex word *prasino'pos*.

Kata'prasinos (=he who is entirely green) is another example of a complex word in Greek. It can be broken down into *ka'ta-* and '*prasinos* (=green).

Words, which are not complex, are called simple words. A few examples of simple words in Greek are: *ho'ris* (=without)*, 'simera* (=today). It is not possible to break down these words into smaller blocks, therefore they are considered simple words.

In this experiment, you will be presented with pairs of complex word and you will be asked to decide which one do you think is MORE complex.

For example, the word *ekatomiri'uhos* (=millionnaire) is very complex. It can be easily broken down into *ekato'mirio* (=million) and *-uhos*. *Ev'nuhos* (=eunuch), however, is not quite so complex. While it is possible to break *ev'nuhos* down into *'evni* and *-uhos* it does not seem completely natural to do so. *Ev'nuhos* is complex, but not as complex as *ekatomiri'uhos*.

Another example of a complex word is *i'porama* (=pad). Even though you may have never heard the word *i'porama* before, you can understand its meaning, because it can be broken down into *i'po-* and *'rama*. *I'polipsi* (=reputation) is also complex – it can be broken down into *i'po-* and *lipsi*. But *i'polipsi* does not seem as complex as *i'porama*. We do not need to break *i'polipsi* into its parts in order to understand its meaning, and, in fact, it seems slightly unnatural to do so.

For each pair of words below, please read both words silently to yourself and then fill in the word you think that is more complex. It is reminded that with the term 'complex' we do not mean the most complicated and/or difficult word of the pair or the word that you are less likely to have encountered before, rather the word which can be broken down to its meaningful blocks easier than its paired one.

It is very important that you provide an answer for every pair, even if you are not certain of your answer. There is no right or wrong answer, just follow your intuition and provide your best guess.

Thank you very much in advance!"