(A)rise and (a)wake
An investigation of two verb pairs

Alexander Lakaw
Abstract:
In this corpus-based study, the two verb pairs arise and rise and awake and wake are investigated. The paper focuses on seven research questions that are related to the meanings of the verbs in question, the semantic specialisations of those verbs, and the semantic relation of the specific verb pair constituents. Furthermore, tendencies of language change are investigated, and an attempt is made to generalise over the influence of the prefix a- on those tendencies.

The results of the quantitative and qualitative analysis show that the verbs awake and wake are more synonymous than the verbs arise and rise. It seems as if due to this difference, the two verbs arise and awake are subject to different processes of language change that take their development into different directions. The observations made about the characteristics of the prefix a- that is involved in the two verbs arise and awake are often ambiguous and inconclusive. Therefore, the influence of this prefix on the processes of language change needs to be analysed further by investigating more word pairs distinguished by the absence or presence of the prefix a-.

Keywords:
corpus linguistics, degrammaticalization, grammaticalization, language change, morphology, prefixation, synonymy, semantic prosody
Table of contents

1. Introduction ........................................................................................................................................... 1
   1.1 Aim and scope ................................................................................................................................... 1

2. Historical background: etymology ......................................................................................................... 3
   2.1 *(A)rise ............................................................................................................................................... 3
   2.2 *(A)wake .......................................................................................................................................... 3
   2.3 The prefix *a*- in *arise* and *awake* ............................................................................................... 4
      2.3.1 Affixation and the prefix *a-*: an attempt to generalise ............................................................... 4
      2.3.2 A characterisation of the prefix *a-* ......................................................................................... 6

3. Theoretical background ........................................................................................................................... 7
   3.1 Corpus linguistics ............................................................................................................................... 8
      3.1.1 Collocation ............................................................................................................................... 9
      3.1.2 Semantic prosody ........................................................................................................................ 10
   3.2 Synonymy ........................................................................................................................................... 10
   3.3 Grammaticalisation ............................................................................................................................ 12
      3.3.1 Definitions and perspectives ....................................................................................................... 12
      3.3.2 Mechanisms ............................................................................................................................... 13
      3.3.3 Unidirectionality and degrammaticalisation ............................................................................. 14

4. Material and method ............................................................................................................................... 15
   4.1 Collection of data ............................................................................................................................... 15
   4.2 Presentation of the material ............................................................................................................... 17
   4.3 Methods ............................................................................................................................................ 18
      4.3.1 Quantitative methods ................................................................................................................ 19
      4.3.2 Qualitative methods .................................................................................................................. 20

5. Quantitative analysis .............................................................................................................................. 21
   5.1 Forms of *(a)rise* and their distribution in the BNC ....................................................................... 22
   5.2 Forms of *(a)rise* and their distribution in the CAE ..................................................................... 24
   5.3 Summary: the verb form analysis of *[ARISE]* and *[RISE]* ....................................................... 25
   5.4 Forms of *(a)wake* and their distribution in the BNC .................................................................. 26
   5.5 Forms of *(a)wake* and their distribution in the CAE .................................................................. 28
   5.6 Summary: the verb form analysis of *[AWAKE]* and *[WAKE]* .................................................. 29
   5.7 Collocations with *(a)rise* ................................................................................................................. 30
      5.7.1 Pre-node collocates of *(a)rise* ............................................................................................... 30
      5.7.2 Post-node collocates of *(a)rise* ............................................................................................ 33
5.8 Collocations with (a)wake ......................................................... 37
  5.8.1 Pre-node collocates of (a)wake ........................................ 37
  5.8.2 Post-node collocates of (a)wake ....................................... 41
5.9 Summary: the collocates of (a)rise and (a)wake ............................ 44

6. Qualitative analysis ........................................................................ 45
  6.1 Meanings of (a)rise .................................................................. 47
    6.1.1 The most frequent meanings of (a)rise ................................. 52
    6.1.2 Transitivity of (a)rise .......................................................... 57
  6.2 Meanings of (a)wake ................................................................. 57
    6.2.1 The most frequent meanings of (a)wake ............................... 59
    6.2.2 Transitivity of (a)wake ....................................................... 60
  6.3 Summary of the qualitative analysis ............................................ 62

7. Discussion .................................................................................. 63
  7.1 Synonymy, semantic specialisation and semantic prosody ............. 64
  7.2 (De)grammaticalisation, lexicalisation and unidirectionality .......... 65
  7.3 The role of the prefix a- ............................................................ 70

8. Conclusion .................................................................................. 72
References ..................................................................................... 76
1. Introduction

This paper is a further development of my bachelor thesis (Lakaw 2007) on the verbs arise and rise. Interesting findings as well as questions that arose during the earlier investigation encouraged further research on the topic of the development of those verbs and motivated me to write this paper. The basic aim of this study is to shed some more light on the matter, and to get a more general picture of the development of verb pairs distinguished by the presence or absence of the prefix a-. To be able to achieve that, the focus had to be broadened: the two verb pairs arise/rise and awake/wake, whose members have similar relations to each other, will be investigated, and a broad range of theories will be applied to look at the use and the meanings of those verbs from various angles.

Compare, for instance, the two imperative sentences Wake up and rise up! and Awake and arise! Both of them predominantly occur in political or religious contexts, and we can easily see their similarity in form and meaning. If both phrases are so similar, then why is the former expression far more common than the latter one (a Google search revealed 12,300 hits for the former and 920 hits for the latter phrase)? Having a look at another two sentences expressing the same meaning, we can see that it seems to be more common to say Wake up and arise! (Google: 3,300) instead of using the similar version Awake and rise up! (Google: 482 hits). Given the similarity of all four sentences, a more even distribution could be expected if they were distributed by chance. But as language use seems to favour some expressions over others, it would be interesting to look into the meanings and use of those four verbs more deeply. Semantic specialisation, bleaching and other processes of language change seem to have an impact on today’s usage of the verbs in question – and in this paper an attempt is made to detect this impact by investigating a large amount of data.

1.1 Aim and scope

The Oxford English Dictionary Online¹ (henceforth OED) lists a large number of meanings that can be expressed with the verbs arise, rise, awake and wake. Some of the meanings listed there are described as being obsolete, archaic or poetic, and this information is helpful to gain a more synchronic view on the usage of the verbs. It does not, however, tell us exactly how those verbs are used today – which of all those meanings listed are the most frequent ones, and which words a present-day speaker would typically combine with the verbs in question.

¹ http://www.oed.com
Therefore, one of the aims of this paper is to investigate the present-day status of the verbs by trying to answer the questions below:

a) What meanings are most frequently expressed by *arise* and *rise* and *awake* and *wake*?

More specifically, it is of interest to find out whether the investigation of the material leads to an answer to the following question:

b) Can any sort of semantic specialisation be detected?

Furthermore, this paper will deal with two more detailed questions which focus on the semantic relations both members of the verb pairs *arise/rise* and *awake/wake* have with each other:

c) Are *arise* and *rise* or *awake* and *wake* synonyms? What type of synonymy is there?

d) Can any other semantic relation between *arise* and *rise* or *awake* and *wake* be detected?

Next to investigating the current status and characteristics of the four verbs, it is the aim of this study to go one step further by asking:

e) Is the assumption (made by the OED), that the verb *arise* is being supplanted by *rise*, detectable in language use? Does the same apply for *awake* and *wake*?

f) If so, what are the reasons for those tendencies?

Finally, one could easily take one more step further and try to generalise over the findings by asking:

g) What role does the prefix *a-* generally play in the tendencies observed? Can we generalise about the role of the prefix *a-?*

To be able to answer the seven research questions above, different theories are applied. Therefore various linguistic topics like semantics, corpus linguistics, language change and grammaticalisation have to be discussed (for more information see Section 3). Investigating the verbs in question from different angles is seen as the key task in this paper. Hence it is important to use suitable methods. It has been decided to use two different approaches: a quantitative and a qualitative one. On the one hand, frequency-counts will be used to gather information about the most frequent collocates the verbs occur with and about what grammatical form (e.g. present, past, perfect, etc.) the verbs are most often used in. On the other hand, detailed semantic analyses of concordance lines – which contain examples of the use of the four verbs in question, will be carried out in the attempt to find answers to the research questions above. For more information on the methods and the material used see Section 4.
2. Historical background: etymology

In this section, a brief overview of the historical developments of the four verbs in question is given. A discussion about the development, status and characteristics of the prefix *a*- is also included here. Most of the information presented in the Sections 2.1, 2.2 and 2.3 was taken from the *Oxford English Dictionary Online* (OED).

2.1 *(A)rise*

The verb *rísan* - the Old English (henceforth OE) ancestor of the verb *rise*, is of Old Teutonic origin and has relatives in almost every Germanic language (cf. Old Frisian *rīsa*, Old Norse and Icelandic *rísa*, and Old Swedish *rīsan*). The earliest records of this verb in textual use have been dated back to the year 1000 by the OED. One might however assume that this verb is older than current evidence can account for. According to the OED, in OE times the simple verb *rísan* was extremely rare; the more common use of the compound form *ārísan* in those times might have resulted in a lack of preserved textual evidence of the verb *rísan* prior to the year 1000.

The origin of the Modern English (henceforth ModE) verb *arise* and its OE ancestor *ārísan* is an Old Teutonic derivation. Its prefix *ā*- adds the meaning ‘up’, ‘out’ and ‘away’ to the stem *rísan*, which itself is already describing an ‘upwards movement’. By the use of the prefix *ā*- the meaning of the stem becomes more intensified. The earliest recordings of *ārísan* date back to the year 825, but the assumption that this verb found its entry into the OE language a couple of centuries earlier than *rísan* is questionable, as it seems unnatural for an affixed form to be older than the stem of the affixed form itself. The popularity of *ārísan* compared to *rísan* in OE times is hardly debatable however. In Middle English (henceforth ME) times however, as the OED tells us, this was about to change due to the influence of the Old Norse variant *rísa*, which was entering language use in the northern parts of Britain. The use of the simpler *rise* became steadily more common. This development went so far that today *rise* has almost supplanted the use of the verb *arise*.

2.2 *(A)wake*

Describing the etymology of ModE *(to) wake* and *(to) awake* proved to be a bit more difficult, since both of them developed from two distinct but synonymous verbs, which coalesced in ME times.

ModE *wake* originates in the strong OE verb *waecnan* (cf. Swedish *vakna*) and in the weak OE verb *waciaen* (cf. Swedish *vaka*). The strong verb is thought to be of Scandinavian origin (probably Old Norse), and it carried the meanings ‘to become awake’ and ‘to come into
being’. The weak verb *wacian* originates in a pre-Teutonic verb and carried the meanings ‘to watch’ and ‘to be or remain awake’. Both verbs have been found in text from the 9th century, and they were used more and more indiscriminately over time, due to their similarity in form and meaning. In ME times the two verbs coalesced and developed the new meaning ‘to rouse from sleep’, which is a frequent meaning of ModE *wake*.

Similarly to the development of *wake*, ModE *awake* derives from two similar but distinct verbs. The strong OE verb Ɨ*wæcnan* and the weak OE verb Ɨ*wacian* are both derivatives of the prefix Ɨ*- (meaning ‘up’, ‘out’ and ‘away’) and the verbs *wæcnan* and *wacian* respectively (the same verbs ModE *wake* originates in, see above). They can be traced back to the 10th century. The verbs were very similar in form and they were both used to express the intransitive meaning ‘to arise or come out of sleep’. Similarly to *wacian* and *wæcnan*, Ɨ*wæcnan*, Ɨ*wæcnan* and Ɨ*wacian* coalesced in ME times and developed the transitive meaning ‘to rouse someone from sleep’, one of the major meanings ModE *awake* possesses (see Section 6.2).

Summarised in simple words one can say that the verbs *arise* and *awake* are derivatives from the verbs *rise* and *wake* respectively. The prefix Ɨ*- seems to play an important role in that process. Therefore, a more detailed discussion about the etymology and the characteristics of this prefix will be presented below.

### 2.3 The prefix Ɨ*- in arise and awake

After having presented the etymology of the two verbs *arise* and *awake* it seems to be helpful to discuss the etymological development and the characteristics of the prefix Ɨ*- which is involved in the derivation processes described above. This matter proved to be more difficult than expected, because the information one can obtain about this prefix often is ambiguous and inconclusive.

#### 2.3.1 Affixation and the prefix Ɨ*-: an attempt to generalise

When speaking about prefixes, we are dealing with the morphological process of *affixation*. According to Matthews (1991), it can be defined by two characteristics:

> Firstly, the form which results from the operation – we may call this the derived form – will consist of the base – the form that the operation applies to – plus an additional morpheme […] . Secondly, the form which is added (the affix) will be constant; it will be the same whatever particular base the operation applies to.  

Matthews (1991: 131)
The second condition stated in the quote above relates to the constant and static form a morpheme has to possess if it is considered to be an affix. One might think that, given the fact that we can combine the prefix \textit{a-} with various bases, the characteristics of this prefix will be the same, unaffected by the base it has been compounded with. But this is not the case.

Bauer and Huddleston describe the prefix \textit{a-} as follows:

The \textit{a-} prefix we are concerned with here has its source in the preposition \textit{on}; a high proportion of the derivatives it forms are themselves prepositions, but there are also adjectives (including some very common ones such as \textit{afraid, asleep, awake}), and a few adverbs (\textit{aloud}). The bases for adjectival derivatives may be nouns (\textit{afoot}), adjectives (\textit{askew}), or verbs (\textit{atremble}). The formation of adjectives in \textit{a-} is still productive, but probably only with verbal bases, in a small number of (often poetic) coinages, such as \textit{aclutter} or \textit{awhir}. \hfill (Bauer & Huddleston 2002: 1706)

Bauer and Huddleston describe the morpheme \textit{a-} as a class-changing prefix, but in our examples above, when \textit{a-} has been combined with the verbal base \textit{rise} or \textit{wake}, the outcome is still a verb. No change of class took place in this process. Furthermore, in the verbs \textit{arise} and \textit{awake}, the prefix describes an upwards movement rather than adding meanings of the preposition \textit{on} to the bases. As one can see, the upwards movement implied in the verbs \textit{arise} and \textit{awake} is not even mentioned in the characterisation of this prefix. Burnley (1992: 295) gives a possible explanation for this lack of information: in the description of the noun \textit{aclivity}, he makes a statement about the prefix used in this word: “[…] the \textit{a-} prefix has the sense ‘up’ in only a few words of OE origin – e.g. \textit{arise, awake}”. This statement allows us to draw two possible conclusions: either the prefix \textit{a-} is used so rarely in the sense of ‘up’ that this meaning is no longer recognised as a semantic feature of this prefix, or we are dealing with two different \textit{a-} prefixes, one of which expresses the meaning ‘on’ described in Bauer’s quote above, whilst the other one is used to describe an upwards movement.

Unfortunately not even etymology can clear things up completely. Classen (1919: 124) describes the prefix \textit{a-} as one of the “principal Old English prefixes surviving in the modern language […], which is of various origins”. This allows us to assume that there are at least two different \textit{a-} prefixes which derived from different etyma. However, the information about this topic provided by the OED is inconclusive. There, six different types of \textit{a-} prefixes are listed, all of which are of different origins, and many of which have influenced each other. For instance, the first type (\textit{a-, prefix\textsuperscript{1}}) is described by the OED as follows: “OE \textit{a-}, […] implying motion onward or away from a position, hence \textit{away, on, up, out}, and thus with verbs of motion adding intensity; as in \textit{a-bide, a-go, a-rise, a-wake}, and many obs. verbs [boldface
added]. According to this entry, arise as well as awake involve this first type of the prefix a-. Regarding the third type of a- prefixes (a-, prefix3), the OED says that it derives from prepositional phrases involving the prepositions on and a. The former of the two prepositions mentioned here – on – seems also to be one of the meanings that can be expressed with the first type of the prefix a- (highlighted in bold above).

Furthermore, in order to describe the origin of the verb awake, the OED states: “As the earliest texts have onweecnan, the a- in later OE was probably = on-, not A-prefix 1”. Bearing in mind that the verb awake already has been given as an example for derivations involving the first type of the prefix a- (highlighted in bold above), it can be concluded that the OED contradicts itself here.

It seems as if the available etymological information cannot really help us with identifying the origin and the real nature of the prefix which is involved in the two verbs arise and awake. We will therefore move on and concentrate on the characteristics that can be identified for the prefix a-.

2.3.2 A characterisation of the prefix a-

The prefix a- is a bound morpheme, it cannot stand alone and it must be combined with at least one other morpheme. It does not carry any syntactic or grammatical information. That means, when adding a- to a base, neither inflectional (i.e. tense) nor morphological (i.e. word-class) changes occur. It seems as if it only contains certain semantic meanings. Bauer (1993) provides a possible explanation: he claims that class-maintaining prefixes are the vast majority in English (1993: 216), and that some of them do not contain more than semantic, or lexical information:

It seems to me to be a very useful approach in the study of word-formation to distinguish between what might be termed the ‘grammatical’ meaning of a word-formation process and the ‘lexical’ meaning […]. Some processes […] have purely grammatical meanings, others, like the prefixation of un-, have only lexical meaning. Probably the majority have both types of meaning. (Bauer 1993: 189)

What Bauer presents here is a description of the process of derivation: the creation of new words by adding affixes to bases (cf. Stockwell 2001: 11). Hence, the prefix a- is a derivational morpheme. As described in Section 2.1, this prefix also works as a semantic intensifier by emphasising the movements expressed by rise and wake. This behaviour can be explained with the help of etymology. Already Classen (1919: 122f) recognised that “[m]any, and possibly all, of these [derivational] affixes were originally independent words”. He
described what Hopper and Traugott (2003) later defined as the process of **morphologisation**: “the fusing of erstwhile independent elements with each other” (2003: 140). Regarding the etymological information given above, they also made an observation that gains importance for the prefix *a*- investigated in this section:

> The beginning of morphologization must be sought in repeated use of syntactic constructions […]. From a historical perspective the relationship between a stem and an affix can only be considered in the context of the phrasal and even higher-level syntax from which they are derived.

(Hopper & Traugott 2003: 141f)

What Hopper and Traugott explained here can be seen as a generalisation about the development of the prefix *a*- from a free lexical morpheme (most likely a preposition) to the derivational affix it is today. During the process of morphologisation, the semantic information possessed by the former preposition became encoded in the prefix that derived from this preposition. These processes often result in content words which “themselves often contain meaningful parts, known as derivational forms […]. Many derivational forms add a meaning component without affecting the category in question […]. Such derivational morphemes are part of the lexicon and can be called ‘lexical derivational morphemes’” (Hopper & Traugott 2003: 4). It is this group of morphemes the *a*- prefix investigated in this study belongs to.

The aim of this section was to give a satisfactory and clear overview of the etymology and the characteristics of the prefix *a*-. Different approaches lead to the recognition of this prefix as a bound, class-maintaining derivational morpheme whose only purpose it is to modify the meaning of the bases it is bound to. Hopper and Traugott (2003: 4) call such morphemes ‘lexical derivational morphemes’. Finally, the process of morphologisation has been discussed, and by that we already slightly touched the field of grammaticalisation which will be one of the topics of the next section.

### 3. Theoretical background

This section presents theoretical background information that is considered to be important for this study. The topics discussed below are corpus linguistics, synonymy, and grammaticalisation.
3.1 Corpus linguistics

Since the investigation of corpus material is an essential part of this study, some words have to be said about the discipline of corpus linguistics. In this section, the aims and characteristics of corpus linguistics are presented briefly. Furthermore, the basic terms collocation and semantic prosody will be defined and discussed in the subsections below.

A corpus, according to Leech (1991: 8), can be defined as “a sufficiently large body of naturally occurring data of the language to be investigated”. The work with electronically stored text collections gained more and more importance within linguistics during the last few decades. Although the ongoing technological development has a big and positive influence on corpus linguistics – huge amounts of data can be stored, processed and investigated – there are critical opinions about this linguistic discipline. Olohan (2004: 14) gives an example of such criticism:

One of Noam Chomsky’s criticisms of corpus linguistics, for example, is that linguistics is about studying language competence, not performance, and that corpus data gives us information about performance and not about competence. (Olohan 2004: 14)

In other words, Chomsky questions the status of corpus studies as a linguistic discipline, due to its lack of significance for the study of language. This criticism has been answered by the findings of other linguists. In a handout for a seminar at Växjö University in Sweden, Michael Stubbs points out that when we are looking at a number of concordance lines (examples of a word in context, aligned by the corpus-software), two different axes can be determined. The horizontal axis of a concordance defines “a single instance of parole: a fragment of a meaningful speech act” (2007: 1). The vertical axis of a concordance then shows “evidence of langue” (ibid.), as it displays repeated formal patterns. Those two axes represent the Saussurean concepts langue and parole in language use, and they can be easily related to the linguistic terms syntagmatic and paradigmatic as well. This contradicts the criticism mentioned above and thereby legitimises corpus studies as a linguistic discipline. Its basic aim is “to reach a better understanding of the workings of human language” (Aijmer & Altenberg 1991: 3). Due to the development of corpus linguistics “into an important framework where description, model-building and practical application prosper side by side” (ibid.), the significance of corpus studies for linguistics cannot really be denied.
3.1.1 Collocation

The term *collocation* is very closely bound to the discipline of corpus linguistics, though its origin lies in the field of phraseology. In the early 1930’s the phraseologist Harold E. Palmer provided a broadly accepted and widely used definition of this term. In his words, collocations are “successions of words which (for various reasons) are best learned as integral wholes” (Palmer 1933: 8). During decades of research in phraseology, and especially in times of computational corpus studies, Palmer’s definition elicited some problems. Researchers were forced to change this rather broad definition according to their needs. Various definitions therefore became established in phraseology and corpus linguistics. The problem of finding a more detailed but still generally applicable definition of the term *collocation* is discussed below. At the end of the discussion the definition used in this study is presented.

Many linguists, like Moon (1998), use Sinclair’s definition of collocations, characterising a collocation as “the occurrence of two or more words within a short space of each other in a text” (Sinclair 1991, quoted in Moon 1998: 26). According to this definition, collocations are any frequent co-occurrences of words. Hence, even accidental co-occurrences of words would qualify for this definition and could therefore be called collocations. For a purely quantitative investigation, this definition of collocation might be suitable enough to be applied. Since this study uses both a quantitative method and a qualitative approach, a more detailed definition of collocation has to be used, which acknowledges some sort of semantic relation between the items. Yule (1996: 122f) explains, for instance, that collocations are co-occurrences of certain items which are somehow lexically connected, and that those lexical connections can differ in various cultures. Hence, collocating items seem to have a semantic relation, and acknowledging that is important for a good and functional definition of the term *collocation*.

In his attempt to find a more suitable definition, compromising between the frequency-based and the semantically-based views on collocations, Siepmann (2005: 437) criticises that “previous definitions […] have relied too heavily on introspection” and too little on the evidence provided by the corpora. In his view, to be able to grasp every linguistic aspect of the concept *collocation*, one has to loosen “the definition of collocation to a considerable extent” (ibid.). He succeeded in finding a generally valid definition of collocation, which, due to its wide applicability, will be used in this study: “a collocation is any holistic lexical, lexico-grammatical or semantic unit normally composed of two or more words which exhibits minimal recurrence within a particular discourse community” (Siepmann 2005: 438). In other words, collocations are two or more words which are lexically, lexico-grammatically or
semantically related, which are recognised as one unit even without the presence of additional context (i.e. holistic) and which occur significantly frequently to be regarded a collocation. The latter condition is related to the minimal recurrence of collocating items. The definition of this concept is far from clear, and it seems to be up to the individual researcher to decide how frequent two co-occurring lexemes have to be in order to be regarded a collocation.

3.1.2 Semantic prosody
Another term that is closely connected to the discipline of corpus linguistics and which will be used in this study is semantic prosody. It describes the phenomenon, “that a given word or phrase may occur most frequently in the context of other words or phrases which are predominantly positive or negative in their evaluative orientation” (Channel 1999: 38), and that, as a result, “the given word takes on an association with the positive or, more usually, the negative, and this association can be exploited by speakers to express evaluative meaning covertly” (ibid.). Additionally, Olohan (2004: 82) emphasises the relation of semantic prosody and corpus linguistics: “Semantic prosody is another aspect of meaning that can be investigated using corpora; indeed it is difficult to investigate it effectively in any other way”. She claims that the positive or negative associations of words can only be observed in large amounts of data. Therefore it will be commented on the semantic prosody of the verbs arise, rise, awake and wake in the quantitative analysis in Section 5.

3.2 Synonymy
This section discusses the semantic problem of synonymy. Synonyms, in simple words, are words that have the same meaning. Although this short definition might sound easy and quite logical, it is not really accepted amongst semanticists. Different approaches have been used to find a more detailed definition of this semantic phenomenon. Similarly to the discussion about collocations above, some views on synonymy are presented here, and finally one definition will be chosen to be used in this study.

One major view on synonymy was presented by Bolinger (1977), who rejects the feature of complete sameness in meaning as a characterisation of synonyms. In his opinion, all sentences that are built up with different forms also have different meanings, as the lexical items themselves contribute to the meaning of the sentence they form. If we narrow our view from the sentence-level to the word-level, we could summarise that words with different forms also have different meanings, a principle which Bolinger (1977: preface, page x) describes with the well known slogan “One form for one meaning, and one meaning for one form”. But would this not mean that there are no synonyms at all? Bolinger does not want to
go that far. In his opinion, synonyms exist, but a pair of synonyms cannot be considered to have exactly the same meanings. Bolinger (1965: 565f) explains that an overlap of semantic ranges is responsible for synonymous meanings of different lexemes. According to that, not the whole semantic range of words can be considered as synonymous, but parts of the semantic range of one word might have the same meaning as parts of the semantic range of another word. If that is the case, those words can be called synonyms.

With Bolinger’s explanation in mind, various linguists have tried to find a suitable definition of synonymy. Yule (1996: 118) defines it as “two or more forms with very closely related meanings, which are often, but not always, intersubstitutable in sentences”. In this definition, Yule acknowledges that synonymy often depends on the context. He exemplifies this with the help of the nouns answer and reply, which in many contexts express synonymous meanings. But in some contexts, they cannot be used interchangeably, as for instance in the sentence She only had one answer/*reply correct on the test.

Cruse (2004) agrees that synonyms do not necessarily have to express the same meaning in every context. He describes three subtypes of synonymy. First of all there are absolute synonyms, which are completely identical in their meaning (Cruse 2004: 154). Examples of this very rarely occurring semantic phenomenon are kick the bucket, die and pass away – all of which express ‘the dying of a living thing’. But when we are forced to use one of those synonyms in order to express our condolences to the bereaved at a funeral, only pass away seems to be a suitable alternative. Hence, even those absolute synonyms differ in their stylistic characteristics and in the appropriateness of their contextual surroundings. It seems as if real absolute synonyms, which can be exchanged in every imaginable context and situation, are very hard to find. Therefore, in this paper, whenever a difference between two items can be identified, they will not be regarded as absolute synonyms.

Cruse’s second subtype is called propositional synonymy (2004: 155). This type is not restricted to single words and their meanings; it rather refers to the meanings of whole phrases or sentences. Whenever items of a sentence can be exchanged paradigmatically without changing the meaning of the sentence, we speak about propositional synonyms. Hence, the nouns violin and fiddle are propositional synonyms in the sentence He is tuning his violin/fiddle. Although both nouns refer to the same musical instrument they cannot be regarded absolute synonyms, due to their different degree of formality.

Cruse mentions a third type of synonymy: near-synonymy. For him, defining this term is difficult, however, because as he explains, “it is not obvious what principle underlies the distinction [between near-synonymy and non-synonymy]” (2004: 156). Apparently, it is up to
the intuition of the speaker to make this distinction. A useful general definition of near-synonyms has been provided by Xiao and McEnery (2006: 108): “[B]y near synonyms we mean lexical pairs that have very similar cognitive or denotational meanings, but which may differ in collocational or prosodic behaviour”. According to this definition, near-synonyms are not collocationally interchangeable; they cannot be freely substituted in an expression. Furthermore, they are restricted to certain stylistic, prosodic, conceptual or contextual conditions. Examples of near-synonyms are the adjectives *pretty* and *handsome*. Although both mean ‘good looking’, they cannot be regarded as absolute synonyms nor as propositional synonyms, because they refer to different concepts. In reference to human beings, *handsome* is bound to *men*, whereas *pretty* is basically used in reference to *women*.

The investigation in Section 6 will attempt to clarify whether there is a semantic relation between *arise* and *rise* and *awake* and *wake* that can be related to the different types of synonymy described above. For that purpose, the distinction between *absolute synonyms*, *propositional synonyms* and *near-synonyms* will be adopted. From the definitions discussed above, Cruse’s (2004: 155f) definitions of absolute and propositional synonyms as well as Xiao and McEnery’s (2006: 108) definition of near-synonyms will be used in this study.

### 3.3 Grammaticalisation

Finally, after presenting the concepts of *collocation*, *semantic prosody* and *synonymy*, the theoretical background section concludes with presenting the linguistic school of grammaticalisation theory. After defining the term *grammaticalisation*, the mechanisms involved will be discussed briefly. This section ends in a discussion about the *unidirectionality* of grammaticalisation and the counter-process *degrammaticalisation*.

#### 3.3.1 Definitions and perspectives

Hopper and Traugott (2003) provide a two-sided definition of the term *grammaticalisation*: first of all, “‘grammaticalization’ refers to that part of the study of language that is concerned with such questions as how lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions” (2003: 1). Furthermore, in addition to the general description of the term, grammaticalisation can also refer to a process “whereby particular items become more grammatical through time” (2003: 2).

The two definitions mentioned above view grammaticalisation from different perspectives. On the one hand, grammaticalisation can be characterised “as primarily a syntactic, discourse pragmatic phenomenon, to be studied from the point of view of fluid
patterns of language use” (ibid.). Hopper and Traugott (2003) describe this as the synchronic perspective on grammaticalisation. However, the chief perspective on grammaticalisation is historical: “From [the historical] perspective, grammaticalization is usually thought of as that subset of linguistic changes whereby a lexical item or construction in certain uses takes on grammatical characteristics, or through which a grammatical item becomes more grammatical” (ibid.).

We can see in the short discussion above that generally, grammaticalisation is related to language change. In grammaticalisation studies, linguists hope to be able to identify certain cross-linguistic pathways of change (compare the title of Fischer et al.’s (2000) book *Pathways of Change – Grammaticalization in English*), along which certain elements of languages develop or have developed over time. Those pathways are called *clines*. The two different perspectives on grammaticalisation can also be applied to clines:

> From a historical perspective, a cline is conceptualized as a natural ‘pathway’ along which forms evolve, a schema which models the development of forms […]. Synchronically a cline can be thought of as a ‘continuum’: an arrangement of forms along an imaginary line at one end of which is a fuller form of some kind, perhaps ‘lexical’, and at the opposite end a compacted and reduced form, perhaps ‘grammatical’. (Hopper & Traugott 2003: 6)

The description of clines as continua relates to a principle that is thought to be an important characteristic of grammaticalisation in general: unidirectionality. Before this principle will be discussed in more detail, a brief overview over the mechanisms underlying grammaticalisation is presented.

### 3.3.2 Mechanisms

Grammaticalisation involves many different mechanisms, four of which are of importance with regard to the present investigation and are presented here. Those four mechanisms are called renewal, analogy, pragmatic enrichment and semantic bleaching.

Often, language change goes hand in hand with a process called renewal. Renewal, or renovation as it is sometimes called, is “the tendency for periphrastic forms to replace morphological ones over time” (Hopper & Traugott 2003: 8). Due to a mind process that is not visible until its consequential results enter language use, old and unclear parts of lexemes (e.g. prefixes) are replaced by newer and clearer elements with similar meanings (e.g. prepositions), resulting in analytic expressions that replace their synthetic counterpart in language use. It is also possible, however, that those synthetic and analytic expressions
coexist and thereby express synonymous meanings (cf. Hopper & Traugott 2003: 7f). Through analogy such changes spread to other words with the same old and unclear element, slowly driving this element out of language use.

Many linguists relate the loss of semantic content directly to the process of grammaticalisation. But the involvement of bleaching in grammaticalisation is questionable:

Perhaps the most damaging evidence against the automatic association of bleaching and sudden emptying of meaning with grammaticalization comes from evidence that later constraints on structure or meaning can only be understood in the light of earlier meanings. In other words, when a form undergoes grammaticalization from a lexical to a grammatical item, some traces of its original lexical meanings tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution. This phenomenon has been called ‘persistence’. (Hopper & Traugott 2003: 96)

Instead of a sudden and total loss of meanings, there appears to be “a balance between loss of older, typically more concrete, meanings, and development of newer, more abstract ones that at a minimum cancel out the loss. Many are the result of pragmatic strengthening, and increase in informativeness with respect to grammatical function” (Hopper & Traugott 2003: 101). Hence, grammaticalisation involves processes like semantic bleaching and pragmatic enrichment at the same time.

3.3.3 Unidirectionality and degrammaticalisation

The theory of grammaticalisation is a highly debated topic. Next to the supporters of grammaticalisation (e.g. Lehmann, Haspelmath), there are some linguists who question the status of grammaticalisation as a theory of language change (cf. Newmeyer, Lass). Their criticism often relates to the hypothesis of unidirectionality, which is presented below.

Lass (2000: 207f) presents the hypothesis of unidirectionality in the form of three defining claims. First of all, grammaticalisation theorists say that language change involves clines, or pathways, which predefine the direction of development of words or morphemes (i.e. lexical items develop into grammatical ones). Secondly, this predefined development is unidirectional. There can never (or only very rarely) be developments which go ‘up the cline’ (for instance to turn a grammatical item into a lexical one). Thirdly, grammaticalisation theorists expect this hypothesis to be of a universal character: “The apparent strength and empirical support for [the two claims described above] enable a universal claim: that all grammatical items in natural languages ultimately derive from lexical items” (Lass 2000: 208). It is this expectance of universality that Lass and other linguists react against. Many grammaticalisation theorists often rely on the ‘theory of rarity’ (Lass 2000: 212) to dismiss
the existence of a counter-development called *degrammaticalisation*, as for instance Lehmann writes:

> [N]o cogent examples of degrammaticalization have been found. This result is important because it allows us to recognize grammaticalization at the synchronic level. Given two variants which are related by the parameters of grammaticalization […] we can always tell which way the grammaticalization goes, or must have gone. (Lehmann 1995: 19)

Criticising the notion of degrammaticalisation, Haspelmath (2004: 36) refers to the same lack of counterexamples when asking the provocative question: “Do the detractors of grammaticalization studies claim that grammaticalization and antigrammaticalization are equally common?” Lass (2000: 212) answers those arguments by criticising grammaticalisation theorists for having not yet dealt with counterexamples properly, except for stating that they are ‘rare’. To end this discussion with a compromise, it shall once again be referred to Hopper and Traugott (2003: 17) who emphasise that the existence of counterexamples should “alert linguists to the need for caution” and remind them that “language change is not subject to exceptionless physical laws, and that diachronic universals, like synchronic ones, are observed tendencies rather than theoretical absolutes” (ibid.). In Section 7, this problem as well as the discussion about unidirectionality and degrammaticalisation will be related to the findings of the present study.

4. Material and method
This section describes and discusses the material used and the methods applied in this study. An attempt to answer the research questions above will be made by examining a large amount of data from various different angles. Two basic methodological approaches are used, a quantitative and a qualitative one. The findings for each verb are compared with one another and related to different theories.

4.1 Collection of data
As with most corpus-related studies, some problems can occur when investigating corpus material. The collection of data for this study needs to be motivated. In order to make a good and successful investigation of corpus material, there are some principles that have to be taken into consideration. Two of those principles are presented by Stubbs (2000: 7):
(1) Reliance on any single corpus is risky. It is best to combine: largish general corpora […], small specialist corpora […], and very large opportunistic text collections […].

(2) To study […] lexical variability, or to study the discourse prosodies around less frequent phrases, much larger text collections may be necessary. (Stubbs 2000: 7)

For this study, a compromising way of data-collection has been found by using two big corpora that each are divided into several sub-corpora. Using those types of corpora has the advantage that one is able to obtain a huge amount of data on the one hand, which on the other hand can be traced back to every single source-text (this might be necessary to obtain a larger co- and context for instance). Using subdivided corpora results in a compromise over the collection of suitable material for frequency-oriented investigations and semantically-based approaches respectively.

The amount of data is another point that should be thought through thoroughly. When using a small amount of corpus data, one is not able to draw far-reaching conclusions. Too many factors come to mind that can disturb the results. Too few examples of the four verbs in question might result in uneven and incomparable numbers. Additionally, the element of chance is much bigger in smaller sets of data than in larger ones and could lead to unwanted misinterpretations. Hence, the data used should not be too small. Using too large amounts of data on the other hand causes a different kind of problem, as investigators are often prone to miss the wood for the trees. When processing a large amount of data, one is likely to miss important information. Hence, the material used in this paper should have certain characteristics. First of all, it has been decided to use corpus material from the two different language varieties British English (henceforth BrE) and American English (henceforth AmE).

One has also to pay attention to the peculiarities of different text genres stored in different subcorpora; the comparability of this data is an important criterion. That is why the corpora used in this study are built up of similar types of subcorpora, each containing material from different language genres (like newspaper language, spoken language, etc.). The corpora used in this study are the British National Corpus (henceforth BNC), accessed through the BYU-interface, and the new Corpus of American English (henceforth CAE).

<table>
<thead>
<tr>
<th>Subcorpora</th>
<th>BNC</th>
<th>CAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td>10.3</td>
<td>76.6</td>
</tr>
<tr>
<td>Fiction</td>
<td>15.9</td>
<td>69.6</td>
</tr>
<tr>
<td>Newspaper</td>
<td>10.5</td>
<td>73.4</td>
</tr>
<tr>
<td>Academic</td>
<td>15.3</td>
<td>73.0</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>44.6</td>
<td>78.1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>96.6</strong></td>
<td><strong>370.7</strong></td>
</tr>
</tbody>
</table>

Table 4.1 on the left describes the structure and size of both corpora in more detail. We can see that the corpora are quite different in size. The fact that the CAE, counting around 371 million words, is four times bigger than the BNC which consists of circa 97
million words is not considered as problematic for this study, since we will be dealing with normalised frequencies rather than with actual numbers. Note that when accessed through the BYU interface, the size of the BNC is smaller than most other sources describe it (100,467,090 words according to the BNC website\(^2\)). On the question why there is such a difference, Mark Davies, the creator of the interface used in this study, responded: “I wondered the same thing. There must be some symbols (&, #, (), etc that they are counting as words, where I don’t. In terms of the actual 4054 texts and the sentences in them, though, they are exactly the same” (personal communication).

One problem that presents itself in Table 4.1 above is the distribution of the material in the different subcorpora. Whereas the subcorpora of the CAE are all rather equal in size, there appears to be an imbalance in the BNC. Here, the miscellaneous subcorpus is far bigger than the other four subcorpora, containing nearly half of the words stored in the BNC in total. Unfortunately nothing can be done about this problem but being aware of it when dealing with the data taken from this subcorpus.

Finally, and in addition to the minor problems discussed above, it has to be mentioned that both corpora are quite different in their general characteristics. The BNC for instance is a static corpus (cf. Olohan 2004: 45), which means that no material will be added to it. The CAE however is a dynamic corpus. Every year, some 20 million words will be added to the corpus. Therefore conclusions can be drawn only about the material that is currently available in the CAE.

### 4.2 Presentation of the material

This section presents the material which has been taken from the two different corpora BNC and CAE. Table 4.2 below summarises the distribution of the lemmas [ARISE], [RISE], [AWAKE] and [WAKE] in both corpora. Those lemmas include all forms in which the verbs in question can occur (i.e. present, present 3rd person, past, participle, -ing). Note that the infrequent and (mostly) informal forms (a)waked and (a)rised are not included in those lemmas and that those forms are not subject to this investigation. Note also that the present (non-3rd person) form of either verb also includes the infinitive and imperative forms (e.g. arise, Arise!, (to) arise), and that specific remarks will be made about this topic whenever necessary.

\(^2\) [http://www.natcorp.ox.ac.uk/corpus/index.xml?style=pdf](http://www.natcorp.ox.ac.uk/corpus/index.xml?style=pdf)
As we can see in Table 4.2, in both corpora, [RISE] is most frequent, counting a total frequency of 63,908 occurrences. This makes [RISE] almost three times as common as the runner up [WAKE], which is narrowly followed by [ARISE]. Both lemmas occur around 22,000 times in total.

Table 4.2. The distribution of the lemmas [ARISE], [RISE], [AWAKE] and [WAKE] in the BNC and the CAE

<table>
<thead>
<tr>
<th>VERB</th>
<th>BNC Frequency</th>
<th>F per 1MW</th>
<th>CAE Frequency</th>
<th>F per 1MW</th>
<th>TOTAL Frequency</th>
<th>F per 1MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ARISE] 9523</td>
<td>98.6</td>
<td>12502</td>
<td>33.7</td>
<td>22025</td>
<td>47.1</td>
<td></td>
</tr>
<tr>
<td>[RISE] 14963</td>
<td>154.9</td>
<td>48945</td>
<td>132.0</td>
<td>63908</td>
<td>136.8</td>
<td></td>
</tr>
<tr>
<td>[AWAKE] 504</td>
<td>5.2</td>
<td>2443</td>
<td>6.6</td>
<td>2947</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>[WAKE] 3790</td>
<td>39.2</td>
<td>19088</td>
<td>51.5</td>
<td>22878</td>
<td>49.0</td>
<td></td>
</tr>
</tbody>
</table>

[AWAKE] however is the most uncommon of the four lemmas, with a total frequency of only 2,947 occurrences.

Without going to deep into details here, a difference between the two language varieties investigated in this study can already be noticed. The verb arise in all its forms seems to be far more common in BrE than in AmE. The BNC reveals a frequency of around 99 occurrences per one million words for the lemma [ARISE]. Compared to the frequency of around 34 occurrences per million words for this lemma in the CAE, we can say that [ARISE] is three times more common in the BrE material. This is a huge difference with regard to the other lemmas which are more or less evenly distributed within the BrE and AmE material (see Section 5.3 for a more detailed discussion about this topic).

4.3 Methods

The methods used in this study are quite different in their nature. A quantitative approach based on frequency counts will be followed by a more detailed qualitative approach, dealing with actual examples of language use (i.e. concordance lines). In order to motivate the setup of the methods, it can be said that many linguists have argued in favour of this two-sided approach. Jan Aarts, for instance, created a model for the description of language use which consists of four different requirements, two of which shall be pointed out here:

1. the model should allow the combination of a quantitative and a qualitative description of data
2. […]
3. the model should allow the description of the full range of varieties, from spontaneous, non edited language use (usually spoken), to non-spontaneous edited language use (usually written or printed)

(Aarts 1999: 6f, quoted from Olohan 2004: 15)
To look at the data from two different angles is one of the key tasks of this study, since the methods aim at different areas of language use. Investigating only one of them does not allow us to draw far-reaching conclusions about the verbs in question. Relying solely on frequency based approaches, for instance, would make it almost impossible to draw conclusions about the semantic values of the verbs in question. Furthermore, the combination of quantitative and qualitative methods seems to be a modern and frequently used approach for successful investigations in corpus linguistics. Olohan (2004: 86) points out that “corpus studies are criticized for focusing excessively on quantitative analysis […]. Thus, a combination of quantitative and qualitative analysis is, in most cases, desirable, particularly if fuller descriptions of linguistic and translational phenomena are to be given and reasons suggested for their occurrence”. Siepmann (2005: 411) agrees. In his view, pure frequency-based investigations can give us information about the “raw material” but no detailed overview about the material’s semantic characteristics. Pure semantic approaches on the other hand are too fragmentary. They “cannot account for all possible cases” (ibid.). His claim that “[t]here has until now been surprisingly little exchange between the two groups” (ibid.) describes the need for a combination of frequency based and semantic approaches in corpus linguistics.

4.3.1 Quantitative methods

In the quantitative analysis, two different characteristics of the four verbs in question were examined. First of all, the data taken from the BNC and the CAE (see Table 4.2 above) were searched for all the grammatical forms the verbs occur in. This information is needed to gain an overview about the form the four verbs arise, rise, awake and wake are mainly, or barely, used in. The results of this analysis will be presented and discussed in Sections 5.1 and 5.2.

Additionally, some information about the most frequent collocates of the four verbs will be given by examining the data by means of search tools available in the interfaces of both corpora. A distinction will be made between pre-node and post-node collocates. This is necessary, since different types of words tend to occur in different positions to the node.

Since investigators are basically interested in collocational co-occurrences which are based on a relation between the collocating items, the accidental co-occurrence of items is a problem. Therefore, the tools used in this quantitative analysis apply different statistical algorithms, which eliminate the co-occurrences that happened by chance (at least to a certain extent). The statistical measure which underlies the search for collocations in this study is called mutual information score and it can be defined as follows:
Mutual Information is a score that relates one word to another by comparing the probability that the two words occur together because they belong together with the probability that their occurrence together is just by chance (McEnery and Wilson 2001: 86). The higher the mutual information score, the stronger the connection between the two words. If, on the other hand, the words co-occur by chance, the mutual information score will be low. (Olohan 2004: 86)

The MI score measures the force of attraction between words and provides usually quite useful information for linguistic investigations like this one. Unfortunately, the MI score is not totally reliable. Infrequent words for instance, which by chance co-occur with other specific items are recognised as collocates, although there might not be any semantic or lexical relation between them. Hence, collocates with a low general frequency have to be examined more cautiously. Although the MI score and other statistical measures for that matter are a bit problematic, they provide good methods for linguists to extract possible collocations from large amounts of data. A more detailed discussion about the pros and contras of MI score and other statistical measures used in linguistics can be found in Bischof (2007).

Next to the problem of accidental co-occurrences, the amount of data has to be taken into consideration. When investigating the collocates of a word, one has to set limits in order to avoid too much data. If the span is big enough, a word might co-occur with any other word a language possesses; hence the search had to be narrowed down a bit. Therefore, this quantitative analysis was restricted to investigate only the most common collocations within the span of three words before (i.e. -3) and three words after the verbs (i.e. +3). This analysis was carried out for each verb, and the results will be presented in Section 5.7 and 5.8.

4.3.2 Qualitative methods

Section 6 will present the results of the qualitative analysis. Similarly to the quantitative approach above, the qualitative analysis is two-sided as well. The result of the first analysis will provide information about how the verbs in question are used and what meanings\(^3\) they basically express. Another aim of this analysis is to compare the meanings of the verbs of each verb pair to see whether there is any semantic relation between them. For this purpose, 250 concordance lines of each verb were randomly chosen and investigated. They were then sorted according to the type of meaning they express. The distribution of the meanings of arise and rise as well as awake and wake were compared to each other and searched for any semantic relation that might exist between those verbs.

\(^3\) Note that in this study no differentiation is made between the terms meaning and sense.
One problem with this type of analysis will be discussed briefly. Channel (1999: 40) quotes Firth, who argues that “the complete meaning of a word is always contextual, and no study of meaning apart from a complete context can be taken seriously”. The topic of the possibility to study meaning is hotly debated in linguistics, and it seems logical that a huge amount of context is needed to understand the complete meaning of a word. The question however is: what can be considered as ‘complete context’? Does ‘complete context’ refer to what is expressed on the phrase-level, sentence-level or do we even need a whole text to understand the complete meaning of a word? It might depend on the actual case how much context is needed to gain an impression of a complete context. Therefore, in the semantic analysis, as much co- and context as needed to understand the complete meaning of the verb in question has been investigated.

The second analysis presented in Section 6 has the purpose to investigate the transitivity of the verbs arise, rise, awake and wake. The same 250 concordance lines of each verb were used to determine whether the verbs in question are bound to direct objects (transitive use) or not bound to direct objects (intransitive use). It might be possible to encounter differences in meaning between the transitive and intransitive use of the verbs.

5. Quantitative analysis

In this section the quantitative analyses are presented. First of all, an analysis of the verb form distribution will be carried out. After that, attention will be paid to the collocates of the verbs in question.

Sections 5.1 and 5.2 describe the verb form distributions of the verbs arise and rise in the BNC and the CAE respectively. Section 5.3 contains a short summary of the findings made in Sections 5.1 and 5.2. The following Sections 5.4 and 5.5 then discuss the verb form distribution of the verbs awake and wake respectively, and findings made in those two sections are summarised in Section 5.6.

It has to be pointed out that, during those investigations, a differentiation has been made between the written and the spoken language material of the BNC and the CAE. Olohan (2004: 46) quotes Kennedy who points out that “we might all use spoken language much more often than written every day, but most instances of written language are likely to have a wider and larger audience than most spoken utterances”. Due to the lack of good and freely available spoken language corpora which are comparable to the material from the written language subcorpora taken from the BNC and CAE, it will not be possible to deal with spoken language in any great detail. However, this study will not exclude the spoken language
variety, since a small amount of spoken language material is present in both the BNC and CAE. But it must be pointed out that the main focus of this investigation is on the written language material, and that no far-reaching conclusions about the verbs in spoken language should be expected.

Sections 5.7 and 5.8 present the collocations of the verbs *arise* and *rise* and *awake* and *wake* respectively. The last subsection 5.9 briefly summarises and discusses the findings related to the investigation of the collocates of the verbs *arise, rise, awake* and *wake*.

### 5.1 Forms of *(a)rise* and their distribution in the BNC

In both corpora, the verbs *arise* and *rise* appear in different forms. Due to the morphological process of *modification*, words are partially or even totally modified in order to express different grammatical meanings (Matthews 1991: 136). Modification includes various different forms, but a very common one is called *vowel change*, with the help of which we are for instance able to express the difference between singular *man* and plural *men* just by altering the vowel. Similarly we are able to express different tenses by altering the stem vowel of the verbs in question. Both *arise* and *rise* follow the same pattern: in order to express the past and the perfect tense, the stem vowel [aɪ] of present tense *(a)rise* changes into [ə] *(a)rose* and [ɪ] *(a)risen*). Additionally, we are dealing with two more verb forms, i.e. the present 3rd person singular *(a)ries* and the –ing form *(a)rising*, both of which are products of a suffixation process.

Tables 5.1.a and 5.1.b display the distribution of the verbs *arise* and *rise* in the BNC while Tables 5.2.a and 5.2.b in the following Section 5.2 summarise the occurrences of *arise* and *rise* in the CAE material. In order to address the variety of forms the verbs can occur in, the term lemma shall be used again.
Table 5.1.a above presents the distribution of the different verb forms of the lemmas [ARISE] and [RISE] in the written language subcorpora of the BNC. There, [ARISE] occurs most often in the subcorpora containing academic and miscellaneous texts. It is most frequent in the academic subcorpus, occurring 245 times per one million words. In the fiction and newspaper subcorpus, this lemma is not very frequent, having a frequency of 15 and 35.1 occurrences per one million words.

Regarding the verb form distribution, it can be noticed that [ARISE] most often occurs as arise. It would be wrong however to claim that this verb is most often used in its present (non-3rd person) form, since an investigation of 100 concordance lines revealed that only 52 percent of this form are actually expressing present tense. 46 percent are infinitive forms, and the remaining 2 percent form imperative expressions. The second most common verb forms [ARISE] occurs in are the –ing form arising and the present 3rd person form arises, resulting in the observation that [ARISE] generally is most favourably used in present tense. The least common form [ARISE] occurs in is arisen, occurring only 7.9 times per one million words in all the written language subcorpora.

Consulting the frequency scores given in Table 5.1.a, we can see that [RISE] is most frequently used in the fiction subcorpus (219.7 instances per 1MW), followed by the newspaper subcorpus (195.8 instances per 1MW). Least common is [RISE] in the academic language genre, although the frequency of 105 instances per one million words can still be regarded as high.

In the written language subcorpora of the BNC, [RISE] is most frequently used in the past form rose, which, with 64.4 instances per one million words, is far more common than the runners up rising and rise. The least common verb form is the present 3rd person form rises. Those findings are in contrast with the observation made above, that [ARISE] is often used in present tense forms, resulting in a noticeable behavioural difference between the two verbs.

Table 5.1.b on the left describes the distribution of the forms of [ARISE] and [RISE] in the spoken language subcorpus of the BNC. We can see that arising is the most common form [ARISE] occurs in, resulting in a frequency of 13 instances per one million words. Since the runners up are

<table>
<thead>
<tr>
<th>Verb form</th>
<th>F</th>
<th>F/1 MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>arise</td>
<td>99</td>
<td>9.6</td>
</tr>
<tr>
<td>arises</td>
<td>50</td>
<td>4.8</td>
</tr>
<tr>
<td>arose</td>
<td>50</td>
<td>4.8</td>
</tr>
<tr>
<td>arisen</td>
<td>28</td>
<td>2.7</td>
</tr>
<tr>
<td>arising</td>
<td>134</td>
<td>13.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>361</td>
<td>34.9</td>
</tr>
<tr>
<td>rise</td>
<td>111</td>
<td>10.7</td>
</tr>
<tr>
<td>rises</td>
<td>31</td>
<td>3.0</td>
</tr>
<tr>
<td>rose</td>
<td>42</td>
<td>4.1</td>
</tr>
<tr>
<td>risen</td>
<td>77</td>
<td>7.4</td>
</tr>
<tr>
<td>rising</td>
<td>96</td>
<td>9.3</td>
</tr>
<tr>
<td>TOTAL</td>
<td>357</td>
<td>34.5</td>
</tr>
</tbody>
</table>
arise, arises and arose, and since the least common verb-form is arisen, it can be concluded that the distribution of the verb forms of [ARISE] in the spoken language material is similar to the pattern observed in the written language subcorpora above. [RISE], however, behaves totally different in the spoken language than it does in the written language material. Most frequently this verb occurs in its uninflected form rise (10.7 instances per 1MW), and according to a brief investigation of 100 concordance lines, in around 64 percent of those instances, the form rise is used as an infinitive, often accompanied by an auxiliary. Rose – being the most frequent form in the written language material – is far behind with a frequency of only 4.1 instances per one million words. Furthermore, it has to be noticed that both verbs are equally frequent in the spoken language subcorpus of the BNC (both verbs having a similar frequency of around 35 instances per 1MW). Compared to the huge difference in the overall frequency of the two verbs displayed in Table 5.1.a above, this observation is remarkable. A satisfying explanation for this observation cannot be found however.

5.2 Forms of (a)rise and their distribution in the CAE

Table 5.2.a below describes the distribution of the verbs arise and rise in the CAE, representing the AmE variety in this study.

<table>
<thead>
<tr>
<th>Verb form</th>
<th>Fiction F</th>
<th>F/1MW</th>
<th>Newspaper F</th>
<th>F/1MW</th>
<th>Academic F</th>
<th>F/1MW</th>
<th>Miscellaneous F</th>
<th>F/1MW</th>
<th>TOTAL F</th>
<th>F/1MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A arise</td>
<td>229</td>
<td>3.3</td>
<td>468</td>
<td>6.4</td>
<td>2591</td>
<td>35.5</td>
<td>1013</td>
<td>13.0</td>
<td>4301</td>
<td>14.6</td>
</tr>
<tr>
<td>A arises</td>
<td>92</td>
<td>1.3</td>
<td>209</td>
<td>2.8</td>
<td>1430</td>
<td>19.6</td>
<td>498</td>
<td>6.4</td>
<td>2229</td>
<td>7.6</td>
</tr>
<tr>
<td>R rose</td>
<td>420</td>
<td>6.0</td>
<td>426</td>
<td>5.8</td>
<td>1314</td>
<td>18.0</td>
<td>697</td>
<td>8.9</td>
<td>2857</td>
<td>9.7</td>
</tr>
<tr>
<td>I arisen</td>
<td>88</td>
<td>1.3</td>
<td>136</td>
<td>1.9</td>
<td>418</td>
<td>5.7</td>
<td>181</td>
<td>2.3</td>
<td>823</td>
<td>2.8</td>
</tr>
<tr>
<td>S arising</td>
<td>81</td>
<td>1.2</td>
<td>134</td>
<td>1.8</td>
<td>1014</td>
<td>13.9</td>
<td>242</td>
<td>3.1</td>
<td>1471</td>
<td>5.0</td>
</tr>
<tr>
<td>E TOTAL</td>
<td>910</td>
<td>13.1</td>
<td>1373</td>
<td>18.7</td>
<td>6767</td>
<td>92.7</td>
<td>2631</td>
<td>33.7</td>
<td>11681</td>
<td>39.7</td>
</tr>
<tr>
<td>R rise</td>
<td>2585</td>
<td>37.1</td>
<td>2544</td>
<td>34.7</td>
<td>1678</td>
<td>23.0</td>
<td>3341</td>
<td>42.8</td>
<td>10148</td>
<td>34.5</td>
</tr>
<tr>
<td>R rises</td>
<td>1833</td>
<td>26.3</td>
<td>684</td>
<td>9.3</td>
<td>535</td>
<td>7.3</td>
<td>1264</td>
<td>16.2</td>
<td>4316</td>
<td>14.7</td>
</tr>
<tr>
<td>I rose</td>
<td>6324</td>
<td>90.9</td>
<td>3303</td>
<td>45.0</td>
<td>1509</td>
<td>20.7</td>
<td>2620</td>
<td>33.6</td>
<td>13756</td>
<td>46.8</td>
</tr>
<tr>
<td>S risen</td>
<td>663</td>
<td>9.5</td>
<td>1112</td>
<td>15.1</td>
<td>636</td>
<td>8.7</td>
<td>1073</td>
<td>13.7</td>
<td>3484</td>
<td>11.8</td>
</tr>
<tr>
<td>E rising</td>
<td>4033</td>
<td>58.0</td>
<td>3325</td>
<td>45.3</td>
<td>1878</td>
<td>25.7</td>
<td>3501</td>
<td>44.8</td>
<td>12737</td>
<td>43.3</td>
</tr>
<tr>
<td>TOTAL</td>
<td>15438</td>
<td>221.8</td>
<td>10968</td>
<td>149.4</td>
<td>6236</td>
<td>85.4</td>
<td>11799</td>
<td>151.1</td>
<td>44441</td>
<td>151.1</td>
</tr>
</tbody>
</table>

In the AmE material, [ARISE] is most common in academic writing and least common in fiction and newspaper language. There is one big noticeable difference between BrE and AmE, namely that [ARISE] in general is far more common in the BNC (with a frequency of 106.2 instances per 1MW) than in the CAE (resulting in a frequency of only 39.7 instances per 1MW). The form arise is the most common form of [ARISE], and it is equally often used to express present tense and as an infinitive form accompanied by an auxiliary (both reached
around 49 percent in a random sample of 100 instances). The least common form is *arisen*. All those findings are quite similar to the observations made in the BNC, except for *arising*, which is less common in the CAE than it is in the BrE material.

In the written language subcorpora of the CAE, [RISE] is most common in fiction (221.8 instances per 1MW), followed by the categories miscellaneous (151.1 instances per 1MW) and newspaper (149.4 instances per 1MW). Similarly to the observation made in the BrE material, [RISE] is less common in academic writing. The most common verb form [RISE] occurs in is *rose* (46.8 instances per 1MW), followed by rising (43.3 instances per 1MW) and *rise* (34.5 instances per 1MW), 83 percent of which are used as infinitive form. The same order can be observed in the BNC material. However, the least frequent verb form in the CAE is *risen* (11.8 instances per 1MW), not *rises*.

Table 5.2.b on the left presents the distribution of the verb forms of [ARISE] and [RISE] in the spoken language subcorpus of the CAE. Similarly to the spoken language subcorpus of the BNC, there are again only a few instances of [ARISE]. We can see that for [ARISE] the most common verb form is *arise*, followed by the 3rd person present form *arises*. As for [RISE], *rise* and *rising* are equally distributed in the spoken language material, each resulting in a frequency of 19 instances per one million words. The past tense form *rose*, most common form in the written language subcorpora, is not very far behind, resulting in a frequency of 12.4 instances per one million words. The least common verb form in the AmE spoken language material is *rises*.

### 5.3 Summary: the verb form analysis of [ARISE] and [RISE]

Summarising the two Sections 5.1 and 5.2, it is noticeable that the frequency of [RISE] in the AmE material is quite equal to the figures in the BNC, whereas [ARISE] is far more common in the BrE material. It is not clear what reasons underlie the prominence of the verb *arise* in BrE. One possible explanation is that *arise* is subject of a process of language change, making this verb less and less frequent in AmE as well as BrE, and that this development has gone further in the AmE variety. This would explain the observation made above, that [ARISE] is less frequent in the CAE material than it is in the BNC. As Mair (1998: 155) points out, “the Americanisation of British English, undisputed fact of recent British linguistic history that it...
is, is rarely a straightforward taking over of American norms and preferences but embedded in a very broad drift in which the two varieties are taken into the same direction at slightly differential speeds”.

The most common language genres are different for both verbs. [ARISE], in both language varieties, is most frequently used in academic writing and least common in fiction, whereas [RISE], in BrE as well as AmE, is mainly used in fiction but rarely in academic language. As far as the spoken language material of both corpora is concerned, one can see that [ARISE] is less common in the spoken language subcorpus of the CAE.

5.4 Forms of (a)wake and their distribution in the BNC

Regarding their verbal paradigms, the verbs awake and wake show similar behaviour as the verbs arise and rise: the vowel [eI] in the present form of (a)wake has to be altered to [@U] to create the past tense form (a)woke and the participle form (a)woken. Additionally, through simple suffixation processes, the two forms (a)wakes and (a)waking are created.

Table 5.4.a describes the distribution of the lemmas [AWAKE] and [WAKE] in the BNC. Similarly to Section 5.1, the investigation of the spoken language material was separated from that of the written language subcorpora, and the results are presented in Tables 5.4.b and 5.5.b. We can see that the verb awake is very infrequent in the BNC material. Only 5.7 instances per one million words were found in the whole written language material. Additionally, most of those instances occurred in the fiction subcorpus, resulting in a frequency of 17.1 instances per one million words. Second most occurrences were found in the miscellaneous subcorpus.

<table>
<thead>
<tr>
<th>Verb form</th>
<th>Distribution of [AWAKE] and [WAKE] in the written language subcorpora of the BNC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>awake 50 3.10 3 0.30 6 0.40 27 0.60 86 1.00</td>
</tr>
<tr>
<td>W</td>
<td>woke 182 11.50 17 1.60 10 0.60 112 2.50 321 3.70</td>
</tr>
<tr>
<td>A</td>
<td>awaken 37 2.30 6 0.60 1 0.10 18 0.40 62 0.70</td>
</tr>
<tr>
<td>K</td>
<td>awaking 2 0.10 0 0.00 1 0.10 2 0.04 5 0.06</td>
</tr>
<tr>
<td>TOTAL</td>
<td>272 17.10 26 2.50 30 2.00 166 3.74 494 5.66</td>
</tr>
<tr>
<td>W</td>
<td>wake 639 40.20 119 11.30 44 2.90 367 8.20 1169 13.50</td>
</tr>
<tr>
<td>A</td>
<td>woke 800 50.30 85 8.10 16 1.00 257 5.80 1158 13.40</td>
</tr>
<tr>
<td>K</td>
<td>woken 238 15.00 47 4.50 9 0.60 135 3.00 429 5.00</td>
</tr>
<tr>
<td>E</td>
<td>waking 200 12.50 30 2.90 15 1.00 175 3.90 420 4.90</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1945 122.30 301 28.70 95 6.20 1018 22.80 3359 38.90</td>
</tr>
</tbody>
</table>
Due to the large size of this corpus however, the 166 occurrences result only in a frequency of 3.7 instances per one million words, a similar frequency as in the subcorpora newspaper and academic, where [AWAKE] is almost non-existent (with a total of 26 and 30 occurrences respectively). Concerning the verb form distribution, the most common form is *awoke*, used in 65% of all occurrences. The least common verb forms are *awaking* and *awakes*, both of which are almost non-existent in the written language subcorpora of the BNC.

Regarding the verb *wake* we can see that it is far more frequent, with an overall frequency of 38.9 instances per one million words. [WAKE] is used most often in fictional writing. This subcorpus contains around 58% of all occurrences of this verb. As concerns the frequency per million words, [WAKE] is second most common in the newspaper category (28.7 instances per 1MW), followed closely by the miscellaneous subcorpus (22.8 instances per 1MW). The verb forms [WAKE] occurs most frequently in are *wake* and *woke*, both having a frequency of around 13.5 instances per one million words. Least common is [WAKE] in the 3rd person present form *wakes*.

### Table 5.4.b. Distribution of [AWAKE] and [WAKE] in the spoken language subcorpus of the BNC

<table>
<thead>
<tr>
<th>Verb form</th>
<th>F</th>
<th>F/1 MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>awake</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>awakes</td>
<td>0</td>
</tr>
<tr>
<td>W</td>
<td>woke</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>awoken</td>
<td>0</td>
</tr>
<tr>
<td>K</td>
<td>awaking</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>TOTAL</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verb form</th>
<th>F</th>
<th>F/1 MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>wake</td>
<td>199</td>
</tr>
<tr>
<td></td>
<td>wakes</td>
<td>44</td>
</tr>
<tr>
<td>A</td>
<td>woke</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>woken</td>
<td>28</td>
</tr>
<tr>
<td>K</td>
<td>waking</td>
<td>37</td>
</tr>
<tr>
<td>E</td>
<td>TOTAL</td>
<td>431</td>
</tr>
</tbody>
</table>

Table 5.4.b on the left describes the distribution of [AWAKE] and [WAKE] in the spoken language subcorpus of the BNC. Within this BrE spoken language material, [AWAKE] is even more uncommon than it is in the written language material. [WAKE] however is quite common even in spoken BrE, resulting in a frequency of 41.7 instances per one million words. In comparison to the written language subcorpora discussed above, after being most common in fiction, [WAKE] is second most common in spoken language. Most frequently, [WAKE] occurs in its uninflected form *wake* and past *woke*, following the same pattern as has been observed in the written language subcorpora. All those observations indicate that [WAKE] is preferred in narrative-like language: fiction, newspaper and spoken language seem to share characteristics (e.g. informal language use, narrative expressions, dialogues) which make these three types of language quite similar. Due to the lack of evidence, this statement should be regarded a hypothesis that needs to be further investigated rather than a fact.

Compared to the findings for [ARISE] above, which in the spoken language subcorpus of the BNC is as common as [RISE], the difference in frequency between the verbs *awake* and
wake is enormous. Since this paper uses only one corpus containing BrE material, drawing far-reaching conclusions is not really possible. It seems however, that [AWAKE] is extremely uncommon in BrE, especially in spoken language.

5.5 Forms of (a)wake and their distribution in the CAE

Tables 5.5.a and 5.5.b describe the distribution of [AWAKE] and [WAKE] in the written language material and the spoken language subcorpus of the CAE.

Table 5.5.a. Distribution of [AWAKE] and [WAKE] in the written language subcorpora of the CAE

<table>
<thead>
<tr>
<th>Verb form</th>
<th>Subcorpus</th>
<th>Fiction F</th>
<th>Fiction F/1MW</th>
<th>Newspaper F</th>
<th>Newspaper F/1MW</th>
<th>Academic F</th>
<th>Academic F/1MW</th>
<th>Miscellaneous F</th>
<th>Miscellaneous F/1MW</th>
<th>TOTAL F</th>
<th>TOTAL F/1MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>A awake</td>
<td></td>
<td>333</td>
<td>4.80</td>
<td>59</td>
<td>0.80</td>
<td>29</td>
<td>0.40</td>
<td>112</td>
<td>1.40</td>
<td>533</td>
<td>1.80</td>
</tr>
<tr>
<td>A awakes</td>
<td></td>
<td>48</td>
<td>0.70</td>
<td>9</td>
<td>0.10</td>
<td>10</td>
<td>0.10</td>
<td>18</td>
<td>0.20</td>
<td>85</td>
<td>0.30</td>
</tr>
<tr>
<td>W awake</td>
<td></td>
<td>1030</td>
<td>14.80</td>
<td>206</td>
<td>2.80</td>
<td>73</td>
<td>1.00</td>
<td>298</td>
<td>3.80</td>
<td>1607</td>
<td>5.50</td>
</tr>
<tr>
<td>A awoken</td>
<td></td>
<td>57</td>
<td>0.80</td>
<td>5</td>
<td>0.10</td>
<td>3</td>
<td>0.04</td>
<td>8</td>
<td>0.10</td>
<td>73</td>
<td>0.20</td>
</tr>
<tr>
<td>K awaking</td>
<td></td>
<td>14</td>
<td>0.20</td>
<td>4</td>
<td>0.10</td>
<td>7</td>
<td>0.10</td>
<td>4</td>
<td>0.10</td>
<td>29</td>
<td>0.10</td>
</tr>
<tr>
<td>E TOTAL</td>
<td></td>
<td>1482</td>
<td>21.30</td>
<td>283</td>
<td>3.90</td>
<td>122</td>
<td>1.64</td>
<td>440</td>
<td>5.60</td>
<td>2327</td>
<td>7.90</td>
</tr>
</tbody>
</table>

Regarding [AWAKE] we can see that in the CAE some more instances of this verb have been found (i.e. 2327 in total). Taking into account the big size of this corpus however, those result in a frequency of only 7.9 instances per one million words, which is quite close to that which has been calculated for the BNC corpus in Table 5.4.a above. It is therefore appropriate to conclude that [AWAKE] is also quite uncommon in the AmE material investigated in this study. Almost 64% of all instances of [AWAKE] occur in the fiction subcorpus (resulting in a frequency of 21.3 instances per 1MW). Similarly to the findings in the written language subcorpora of the BNC, [AWAKE] seems to be almost exclusively used in fictional writing, since only small numbers of occurrences can be found in the remaining categories newspaper, miscellaneous and academic, the latter of which contains least occurrences of this verb. The most frequent verb form [AWAKE] is used in the AmE material is past awake. Almost 70% of all instances of [AWAKE] are in the past tense form. Next most frequent is the form awake, reaching 23% of all occurrences, 50 percent of which express present tense, while in 40 percent of the cases, arise was used as an infinitive. In the remaining 10 percent, awake is used in imperative form. Similarly to the observations in the BNC, the forms awakes, awoken and awaking are almost non-existent.
In the CAE, [WAKE] is even more frequent, reaching a total frequency of 54.1 instances per one million words. Similarly to the findings in the BNC material, most occurrences were found in the fiction subcorpus (containing 62% of all occurrences). A difference, however, is that [WAKE] in the runner up categories miscellaneous and newspaper is more frequent than in the BrE material. In the AmE material, the verb *wake* is very uncommon in academic writing. This observation is an additional support for the hypothesis that [WAKE] has a tendency to be favoured in narrative-like language. The most frequent verb form [WAKE] occurs in is uninflected *wake* (22.5 instances per 1MW) and past tense *woke* (18.3 instances per 1MW), the same verb forms that are most frequent in the BrE material.

Table 5.5.b on the left describes the distribution of [AWAKE] and [WAKE] in the spoken language subcorpus of the CAE. [AWAKE] is once again only sparsely present in the spoken language material, 116 occurrences were found. Most of them, almost 53% of all instances, were the past tense form *awoke*. Almost 39% of all instances appeared in the form *awake*. For [AWAKE], those two verb forms were most frequent in the written language and spoken language material of both the BNC and CAE.

The frequency for [WAKE] displayed in Table 5.5.b above is identical to that present in Table 5.4.b: 41.7 instances per one million words can be found in the spoken language subcorpora of both the BNC and the CAE. This means that yet again, in comparison with the different subcorpora containing written language material of AmE, [WAKE] is second most frequent in the spoken language subcorpus. Similarly to the observations made above, the most common form for this verb is *wake*, followed by past tense *woke*.

### 5.6 Summary: the verb form analysis of [AWAKE] and [WAKE]

This section summarises the observations made in Section 5.4 and 5.5 above. Regarding the overall frequencies of both verbs in the BrE and AmE material, [WAKE] is far more common than [AWAKE] in both language varieties. In both the BNC and the CAE, the verb *wake* in all its forms is seven times more frequent than *awake*. It can be pointed out, that [WAKE] is somewhat more frequent in the AmE material, whereas [AWAKE] is almost equally uncommon in both corpora. In contrast to the observations in Section 5.1 and 5.2, both verbs investigated in the former two sections are favoured in the same language genre. In both
language varieties, [AWAKE] and [WAKE] are most frequently used in the fiction category. Furthermore, [WAKE] as well as [AWAKE] are most often used in their past tense forms (a)woke and their uninflected form (a)wake. We can also see that other verb forms of [AWAKE] are only sparsely present in both corpora. This lack of variation could be due to the low frequency of this verb in both the AmE and the BrE variety. It could also be the case however, that both the simple present and the simple past in general are the most frequent forms in which verbs are used, and that the figures for [AWAKE] result from a general tendency of speakers to prefer those two verb forms over others.

Concerning the spoken language corpora of the BNC and the CAE, differences could be detected. The observations that [AWAKE] is almost non-existent in both spoken language subcorpora investigated, and that [WAKE] is very common in those might indicate a certain stylistic specialisation of the two verbs. Next to the observation that [WAKE] is very common in spoken language, it was found that in written language [WAKE] is mainly used in the fiction subcorpora. This latter observation supports the thesis that the verb wake has a tendency to be favoured in narrative-like language, a phenomenon which might be expressed in terms of semantic specialisation. This topic will be discussed in more detail in Section 5.8.

5.7 Collocations with (a)rise

This section presents the most frequent collocates of the verbs arise and rise. Sections 5.7.1 and 5.7.2 discuss the most common collocates in pre-node and post-node position respectively. Various tables are used to present the 10 most frequent collocates, and information is given on frequencies and mutual information score (see Section 4.3.1). The percentage column in the tables below shows how many of the total occurrences of this word in the corpus co-occurred with the verb in question.

5.7.1 Pre-node collocates of (a)rise

Table 5.7.a below presents the top ten pre-node collocates of [ARISE] in the BNC and CAE. The first thing that has to be noticed is that all the listed collocates are nouns, most of which are probably used as subjects. Looking at the left side of Table 5.7.a, we can see that seven of the top ten collocates for [ARISE] in the BNC material are words with negative connotations, such as misunderstandings, complications, dispute(s), complication, difficulties, and odours, the former of which has the strongest relation to [ARISE].
Table 5.7.a. Top 10 pre-node collocates of [ARISE] in the BNC and CAE

<table>
<thead>
<tr>
<th></th>
<th>collocate</th>
<th>BNC</th>
<th></th>
<th>collocate</th>
<th>CAE</th>
<th></th>
<th>MI</th>
<th></th>
<th>MI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>misunderstandings</td>
<td>12</td>
<td>5.61</td>
<td>5.24</td>
<td>complications</td>
<td>79</td>
<td>2.11</td>
<td>5.34</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>vacancy</td>
<td>19</td>
<td>5.60</td>
<td>5.24</td>
<td>conflicts</td>
<td>159</td>
<td>1.94</td>
<td>5.26</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>odours</td>
<td>11</td>
<td>4.33</td>
<td>4.98</td>
<td>tumors</td>
<td>35</td>
<td>1.54</td>
<td>5.02</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>complications</td>
<td>36</td>
<td>4.30</td>
<td>4.97</td>
<td>disagreements</td>
<td>23</td>
<td>1.46</td>
<td>4.97</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>disputes</td>
<td>57</td>
<td>3.68</td>
<td>4.82</td>
<td>dilemmas</td>
<td>21</td>
<td>1.45</td>
<td>4.97</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>vacancies</td>
<td>15</td>
<td>3.24</td>
<td>4.69</td>
<td>controversies</td>
<td>16</td>
<td>1.39</td>
<td>4.92</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>tentacles</td>
<td>12</td>
<td>2.88</td>
<td>4.57</td>
<td>disputes</td>
<td>58</td>
<td>1.27</td>
<td>4.83</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>complication</td>
<td>10</td>
<td>2.79</td>
<td>4.54</td>
<td>difficulties</td>
<td>98</td>
<td>1.10</td>
<td>4.69</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>difficulties</td>
<td>181</td>
<td>2.67</td>
<td>4.50</td>
<td>emergencies</td>
<td>14</td>
<td>1.02</td>
<td>4.62</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>dispute</td>
<td>82</td>
<td>2.54</td>
<td>4.45</td>
<td>tensions</td>
<td>43</td>
<td>0.91</td>
<td>4.50</td>
<td></td>
</tr>
</tbody>
</table>

This can be seen in the high MI score of 5.24 and the high percentage, expressing that 5.61 percent of all instances of *misunderstandings* in the BNC occur within the context of [ARISE]. We can see an even clearer tendency in the CAE, where all top ten collocates have a negative connotation. This observation allows the conclusion that [ARISE] preferably combines with negative words, forming sentences like (001) and (002)

(001) The accused would not be present, and *misunderstandings* sometimes arose between judge and counsel […]. (BNC misc.)

(002) Similar complications arise in the recycling of plastics. (CAE acad.)

The BNC is the only of the two corpora which shows two positive collocates of [ARISE], the singular and plural form of the noun *vacancy*, both of which have a very high MI score of 5.24 and 4.69 respectively. One might argue that in BrE the verb *arise* is bound to negative words to a lesser extent than in AmE. One could also argue that the phrase *vacancies arose* or similar gained a status of a fixed phrase only in BrE, making it frequently used only in this language variety. An argument for the latter is that, in the AmE material, the singular form *vacancy* occurs only 5 times accompanied by the verb *arise*. The plural form *vacancies* in the context of *arise* is not even present in the CAE.

On the other hand, the noun *tumors* seems to be closely related to the verb *arise* only in the AmE variety (due to the high MI score of 5.02), resulting in a fixed phrase *tumors arose* or similar:

(003) Granular cell *tumors* can arise at any anatomic site […]. (CAE acad.)

---

4 Note that the boldface type was added in all examples.
In the BrE material, the term *tumour* can only be found 6 times in the context of *arise*, the plural form is not even present in this context. Another observation with regard to the collocate *tumor* is that it – together with *tentacle* in the BrE material (a neutral noun that can occur in negative circumstances) – represents the only concrete nouns in the lists above. Hence, [ARISE] seems to favour abstract nouns as collocates.

Table 5.7.b contains the 10 most frequent pre-node collocates of [RISE]. Similarly to the observations for [ARISE] above, all of the collocates of the verb *rise* are nouns or noun-phrase (NP) constituents.

Those NP constituents are for instance words like *mid-term, pre-tax* or *second-quarter*, most of which could be falsely identified as parts of phrases where *rise* functions as a noun. Here however, *rise* does function as verb, as the examples below show:

(004) Mid-term net profit rose 107% to $ 608.0m […]. (BNC misc.)
(005) Pre-tax profits rose 14.5 per cent to £39.2 million […]. (BNC news.)
(006) Polaroid said its second-quarter earnings rose 4 percent to $ 29.2 million […]. (CAE news.)

The noun *hackles* has a very strong connection to the verb *rise* in the BNC and the CAE, having a high MI score of 6.65 and 5.79 respectively. The fact that 36.21 percent of all instances of *hackles* in the BNC and 12.9 percent in the CAE occur with [RISE] allows us to conclude that *hackles* is a closely related collocate of this verb. It refers to hairs or feathers of animals, which can be raised as response to certain stimuli. Furthermore, *hackles* is one of four concrete nouns among the top ten collocates of the BNC, leaving six abstract nouns (e.g. *earnings, incomes*) as collocates for [RISE]. In the CAE, there are six concrete nouns, and only four abstract nouns listed as collocates. It seems as if the verb *rise* neither favours concrete or abstract nouns, but is combinable with both types.
In the BNC, six of the top ten collocates are somehow related to the semantic field of ‘economy’ (i.e. turnover, mid-term, earnings, pre-tax, prices, incomes). Although only four of the CAE’s top ten collocates of [RISE] are related to economics, it becomes obvious that the verb rise is closely related to the topics of ‘economy’ and ‘money’.

Regarding the concept of semantic prosody, the collocates of rise in both corpora are mainly neutral or even positive in their connotation, but two negative collocates could be found in the CAE material: floodwaters and bile (a secretion of the gall-bladder, often rising in the oesophagus when one is not feeling very well), both of which have a high MI score of 4.88 and are therefore closely bound to [RISE]:

(007) However, when floodwaters rise, it’s the faithful who rise to the occasion. (CAE misc.)
(008) At the thought of vodka’s sickly tastelessness, bile rose in his throat. (CAE fic.)

However, these observations are not enough to allow the claim that rise is closely connected to either positive or negative words.

5.7.2 Post-node collocates of (a)rise

After having discussed the most frequent collocates in pre-node position, the focus will be on frequently co-occurring words that appear after the node. Table 5.7.c below presents the results for the material taken from the BNC and the CAE.

<table>
<thead>
<tr>
<th>#</th>
<th>collocate</th>
<th>BNC F</th>
<th>%</th>
<th>MI</th>
<th>collocate</th>
<th>CAE F</th>
<th>%</th>
<th>MI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>spontaneously</td>
<td>16</td>
<td>3.61</td>
<td>4.80</td>
<td>spontaneously</td>
<td>34</td>
<td>2.99</td>
<td>5.69</td>
</tr>
<tr>
<td>2</td>
<td>in connection with</td>
<td>56</td>
<td>3.57</td>
<td>4.79</td>
<td>ashes</td>
<td>10</td>
<td>0.40</td>
<td>3.67</td>
</tr>
<tr>
<td>3</td>
<td>as to</td>
<td>97</td>
<td>1.35</td>
<td>3.82</td>
<td>interactions</td>
<td>24</td>
<td>0.36</td>
<td>3.56</td>
</tr>
<tr>
<td>4</td>
<td>indirectly</td>
<td>13</td>
<td>1.25</td>
<td>3.74</td>
<td>concerning</td>
<td>33</td>
<td>0.35</td>
<td>3.53</td>
</tr>
<tr>
<td>5</td>
<td>out of</td>
<td>543</td>
<td>1.16</td>
<td>3.66</td>
<td>independently</td>
<td>10</td>
<td>0.26</td>
<td>3.25</td>
</tr>
<tr>
<td>6</td>
<td>naturally</td>
<td>38</td>
<td>0.92</td>
<td>3.43</td>
<td>from</td>
<td>3340</td>
<td>0.22</td>
<td>3.09</td>
</tr>
<tr>
<td>7</td>
<td>from</td>
<td>2600</td>
<td>0.64</td>
<td>3.08</td>
<td>naturally</td>
<td>20</td>
<td>0.19</td>
<td>2.94</td>
</tr>
<tr>
<td>8</td>
<td>in part</td>
<td>13</td>
<td>0.63</td>
<td>3.06</td>
<td>context</td>
<td>40</td>
<td>0.17</td>
<td>2.85</td>
</tr>
<tr>
<td>9</td>
<td>in respect of</td>
<td>18</td>
<td>0.60</td>
<td>3.01</td>
<td>regarding</td>
<td>29</td>
<td>0.17</td>
<td>2.82</td>
</tr>
<tr>
<td>10</td>
<td>in response to</td>
<td>12</td>
<td>0.60</td>
<td>3.01</td>
<td>primarily</td>
<td>27</td>
<td>0.17</td>
<td>2.82</td>
</tr>
</tbody>
</table>

We can see that the verb arise collocates with two different kinds of items in post-node position. On the one hand, in both corpora, [ARISE] quite often collocates with prepositions or prepositional phrases. Those are for instance out of, in part, in respect of, in response to, in connection with, and from, the latter of which shows a close relation to the verb arise in both corpora, given that the MI score reaches 3.08 in the BNC and 3.09 in the CAE respectively. In
combination with those collocations, *arise* is used in a causative sense. This means that reasons are presented that give answer to the question why something arises:

(009) My position really is that the trade friction **arises from** problems in particular industries. (CAE spok.)
(010) Yet there is in him something that neither **arises out of** culture nor contributes directly to it. (BNC misc.)

On the other hand, [ARISE] frequently co-occurs with adverbs. As can be seen in Table 5.7.c above, the most frequent post-node collocate of the verb *arise* in both corpora is *spontaneously*, reaching high MI scores of 4.8 and 5.69 respectively. Other adverbs present in the top ten lists above are *indirectly, naturally, independently* and *primarily*. In combination with those adverbs, [ARISE] can be used to describe the manner, or how something arises:

(011) Part of the presenter's role is to watch for good photo opportunities as they **arise spontaneously** in the course of events […]. (BNC misc.)
(012) The subject seemed to **arise naturally** out of the last tale. (CAE fic.)

In the top ten list of post-node collocates in the CAE, three nouns can be found, all of which might be part of fixed phrases involving [ARISE]. First of all the noun *ashes* can be found in second rank, with a relatively high MI score of 3.67. This noun is part of the phrase *to arise from the ashes* or similar, a saying that probably traces back to the antique myth of the resurrection of the phoenix:

(013) The **phoenix arose from the ashes**. (CAE spok.)
(014) The Santa Barbara, California-based company **arose from the ashes** of Santa Monica-based National Medical Enterprises […]. (CAE acad.)

We can see in the examples above that the phrase [ARISE] *from the ashes* is no longer solely bound to the word *phoenix*. In fact, only 3 of the 10 instances involving the collocate *ashes* also contained the word *phoenix*. This means that the phrase [ARISE] *from the ashes* has developed additional semantic values, which makes it applicable to other concepts as well.

The second noun in Table 5.7.c above is *interactions*, and it is almost equally strongly bound to *arise as ashes*, according to the MI score of 3.56. This noun is used in the frequent fixed phrase [ARISE] *from interactions*, a causative expression as the example below shows:
In contrast, diabetes arises from interactions between several different genes, spread over several chromosomes. (CAE misc.)

The third and last noun that can be found among the top ten collocates of [ARISE] in the CAE is context. The phrase [ARISE] in the context of can also be regarded as a fixed phrase, which is not as frequent as the former two discussed above. It can be discussed whether this phrase is causative or descriptive, it seems to possess elements of both:

Hakim’s statement, it must be remembered, arose in the context of a discussion about equality. (CAE acad.)

On the one hand, the example above shows that Hakim’s statement somehow originates in a discussion about equality, hence it describes a causative use of the verb arise. On the other hand however, Hakim’s statement also is related to the discussion about equality, a feature that might be more descriptive in a sense.

Table 5.7.d below describes the top ten most frequent post-node collocates of [RISE] in the BrE and AmE material. It can be seen that the use of the verb rise is more uniform than that of arise.

<table>
<thead>
<tr>
<th>#</th>
<th>BNC collocate</th>
<th>F</th>
<th>%</th>
<th>MI</th>
<th>CAE collocate</th>
<th>F</th>
<th>%</th>
<th>MI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>steeply</td>
<td>64</td>
<td>19.39</td>
<td>6.03</td>
<td>majestically</td>
<td>33</td>
<td>15.57</td>
<td>5.97</td>
</tr>
<tr>
<td>2</td>
<td>crescendo</td>
<td>11</td>
<td>8.66</td>
<td>5.22</td>
<td>unsteadily</td>
<td>33</td>
<td>10.96</td>
<td>5.62</td>
</tr>
<tr>
<td>3</td>
<td>sharply</td>
<td>132</td>
<td>5.63</td>
<td>4.79</td>
<td>steeply</td>
<td>54</td>
<td>9.93</td>
<td>5.52</td>
</tr>
<tr>
<td>4</td>
<td>steadily</td>
<td>84</td>
<td>5.20</td>
<td>4.71</td>
<td>prominence</td>
<td>127</td>
<td>7.03</td>
<td>5.18</td>
</tr>
<tr>
<td>5</td>
<td>bait</td>
<td>28</td>
<td>5.11</td>
<td>4.70</td>
<td>steadily</td>
<td>231</td>
<td>4.87</td>
<td>4.81</td>
</tr>
<tr>
<td>6</td>
<td>prominence</td>
<td>29</td>
<td>4.80</td>
<td>4.63</td>
<td>exponentially</td>
<td>28</td>
<td>4.86</td>
<td>4.81</td>
</tr>
<tr>
<td>7</td>
<td>dramatically</td>
<td>71</td>
<td>4.69</td>
<td>4.61</td>
<td>sharply</td>
<td>294</td>
<td>4.60</td>
<td>4.75</td>
</tr>
<tr>
<td>8</td>
<td>ashes</td>
<td>24</td>
<td>4.60</td>
<td>4.59</td>
<td>precipitously</td>
<td>18</td>
<td>4.40</td>
<td>4.71</td>
</tr>
<tr>
<td>9</td>
<td>vertically</td>
<td>16</td>
<td>3.98</td>
<td>4.45</td>
<td>crescendo</td>
<td>18</td>
<td>3.95</td>
<td>4.60</td>
</tr>
<tr>
<td>10</td>
<td>cents</td>
<td>10</td>
<td>3.61</td>
<td>4.35</td>
<td>ashes</td>
<td>89</td>
<td>3.53</td>
<td>4.49</td>
</tr>
</tbody>
</table>

The presence of many adverbs in the table above shows that the basic use of [RISE] is descriptive. Five of the ten collocates in the BNC are adverbs (e.g. sharply, steadily, vertically), the most prominent of which – steeply – is very closely bound to the verb rise, as can be seen in the high MI score of 6.03 and the fact that 19 percent of all instances of steeply in the BNC occurred in the context of [RISE]. Amongst the top ten collocates in the CAE there are even more adverbs. The adverbs majestically, unsteadily, and steeply can be found in the top three ranks, having high MI scores of around 6 to 5.5. It seems as if the verb rise is
more likely to co-occur with adverbs as post-node collocates than it is the case for arise. Another difference between [ARISE] and [RISE] is that the latter is not frequently used in causative senses, but almost exclusively to explain how something rises, not why:

(017) On one side of the road, thickets and small trees rose steeply […]. (BNC fic.)
(018) The moose rises majestically as if posing for a postcard. (CAE spok.)

Three nouns can be seen in Table 5.7.d above that occur in both the BrE and AmE material. First of all, the noun crescendo seems to have a strong connection to [RISE], having high MI scores in both the BNC and the CAE (i.e. 5.22 and 4.6 respectively). In the BNC, 8.66 percent of all instances of crescendo occurred in the context of [RISE], a number which suggests a strong relation between those two words. In the CAE, this percentage is not as high as in the BNC (i.e. 3.95%) but there are more co-occurrences of crescendo and [RISE]. In most of the instances rise and crescendo co-occurred in the fixed phrase [RISE] to a crescendo, only some instances of [RISE] in a crescendo could be found. Both variants seem to have the same meaning however:

(019) Linda's cries rise to a crescendo. (CAE fic.)
(020) Francis sobbed, and Elaine's cries rose in a crescendo […]. (BNC fic.)

The second noun that frequently co-occurs with [RISE] in both corpora is prominence, which has high MI scores in the BNC and the CAE (4.63 and 5.18 respectively). Together, both words form the fixed phrase [RISE] to prominence, which means 'to become famous'.

(021) Alan Travers had risen to prominence in his wife's organisation […]. (BNC acad.)

The third noun that is listed as a frequent post-node collocate of [RISE] in the BrE as well as in the AmE material is ashes. It is part of the fixed phrase [RISE] from the ashes, a synonym of [ARISE] from the ashes, a phrase which was already discussed above. Comparing the frequency of ashes as post-node collocate of arise and rise, we can see that it more frequently co-occurs with the latter verb. Hence, it can be concluded that the phrase [RISE] from the ashes is more frequent than the synonymous expression containing [ARISE].
ONE of Anglesey’s oldest town halls is set to **rise from the ashes** of a fire which destroyed it.

(BNC news.)

She’s like the **phoenix** who **rises from the ashes**. (CAE spok.)

The examples above show two different uses of the phrase. First of all, it seems to be common to connect *[RISE] from the ashes* with the concept of ‘fire’, giving the phrase a somewhat literal sense, although it mainly has to be regarded a figurative expression nonetheless. The second example shows the connection of this phrase to the mystical bird phoenix, referring to the origin of this phrase, which is nowadays used in various circumstances and contexts. The word *phoenix* as pre-node collocate is quite frequent: in the context of *[RISE]* it occurred eight times in the BNC, and as much as 37 times in the CAE.

The last two post-node collocates of *[RISE]* that will be discussed here are the nouns *bait* and *cents*, both of which only occurred in the top ten list of the BNC. The former noun reaches a MI score of 4.7, and it forms the fixed phrase *[RISE] to the bait*, which is a figurative expression with the meaning ‘to react to a temptation or provocation’. The second noun is *cents* and it is another collocate of *[RISE]* that is related to the semantic fields of ‘economy’ and ‘money’, both of which have already been identified as the semantic fields most of the pre-node collocates of *[RISE]* belong to.

Matthew was not going to **rise to the bait** his mother dangled so expertly in front of him.

(BNC fic.)

The pound **rose** almost 2 **cents** to $ 1.4945 […]. (BNC newsp.)

### 5.8 Collocations with *(a)wake*

This section presents the most frequent collocates of the verbs *awake* and *wake*. Sections 5.8.1 and 5.8.2 discuss the most common collocates in pre-node and post-node position respectively. The tables below present the 10 most frequent collocates in the same way as the tables used in Section 5.7.

#### 5.8.1 Pre-node collocates of *(a)wake*

Table 5.8.a below presents the top ten pre-node collocates of *[AWAKE]* and *[WAKE]* in the BNC and the CAE. Regarding the top ten collocates of the BNC, one problem is obvious. Due to the scarcity of instances of *[AWAKE]*, especially in the BNC (see Section 4.2), only a few items in the list can be regarded as real collocates of the verb *awake*. We can for instance see that many pronouns are listed as co-occurring items.
The conclusion that those pronouns are also collocates of [AWAKE] would be wrong however, since pronouns in general are very frequent; they simply are needed to form grammatical sentences. Hence, the occurrence of those pronouns (as well as the verbs was, had and the conjunction but for that matter) is due to the general infrequency of the verb awake, which results in a lack of other items that it would collocate with. The low MI scores of the co-occurring pronouns allow the conclusion that [AWAKE] has no real semantic relation with those words.

Two words in the BNC can be regarded as collocates of [AWAKE] however: the noun morning and the conjunction when, both of which are also listed as collocates in the CAE. There, similar words like dawn, night, and later can be found, all of which share their relation to the concept of ‘time’. Hence, [AWAKE] seems to be preferred in temporal expressions like those in the examples below.

(026) The next morning I awoke early and tidied myself for school. (BNC misc.)
(027) When he awoke again it was seven. (BNC fic.)
(028) Hours later, I awoke on the floor of my lab. (CAE fic.)

Example (026) illustrates the use of morning in the context of [AWAKE]. The noun is the most frequent collocate of awake in the BNC, reaching a MI score of 4.16. In the CAE, morning is a frequent collocate of awake as well, reaching a MI score of 3.95, ending up in third rank of the top ten list. Given its relation to the semantic field of ‘sleep’, it feels natural that [AWAKE] has a semantic connection to expressions like morning, dawn, when and later.

Even words belonging to the semantic field of ‘sleep’ can be found among the top ten collocations of awake, at least in the AmE material. Here, antonymic expressions are quite frequent that involve for instance the words asleep (adj.) and sleep (noun). This is not really
unexpected; research has shown “that antonyms co-occur in sentences significantly more often than chance would predict” (Paradis et al. 2006: 1):

(029)  I fell asleep and awoke. (CAE fic.)
(030)  I fell into a deep sleep and awoke to the shrill ringing of the doorbell. (CAE fic.)
(031)  At 4:00 I go below and immediately fall asleep. I awake at 7:40 in my berth.
       (CAE misc.)

While (029) and (030) show the pairing of the antonymous terms asleep/sleep and awoke within one sentence, the latter example (031) shows that the collocational relation between these antonyms even goes beyond the boundaries of punctuation marks. Although we are able to see that [AWAKE] tends to be used in antonymic expressions, too few examples were found to safely conclude that phrases like in the three examples above represent typical or even frequent uses of the verb awake.

Table 5.8.b below contains the top ten pre-node collocates of [WAKE] in the BNC and the CAE. Compared to the observations made above, the verb wake prefers the company of antonymic expressions to a larger extent than [AWAKE]. Five of the top ten collocates in the BNC are antonymic expressions from the semantic field of ‘sleep’, i.e. slept, asleep, sleep (noun), and dream(s). As many as six words from the same semantic field can be found in the listing for the CAE: dozed, sleeper, asleep, slept, nap, and sleep (noun). Instead of only favouring some specific terms related to the concept of ‘sleep’, [WAKE] seems to be frequently combined with a wide range of terms related to this semantic field. The two antonymic collocates asleep and sleep were also found in Table 5.8.a above and they were already discussed with regard to the collocates of [AWAKE].

Table 5.8.b. Top 10 pre-node collocates of [WAKE] in the BNC and CAE

<table>
<thead>
<tr>
<th>#</th>
<th>collocate</th>
<th>BNC</th>
<th></th>
<th></th>
<th>collocate</th>
<th>CAE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>slept</td>
<td>17</td>
<td>1.04</td>
<td>4.48</td>
<td>dozed</td>
<td>16</td>
<td>3.40</td>
<td>5.39</td>
</tr>
<tr>
<td>2</td>
<td>asleep</td>
<td>19</td>
<td>0.84</td>
<td>4.27</td>
<td>stirs [v.]</td>
<td>12</td>
<td>1.34</td>
<td>4.46</td>
</tr>
<tr>
<td>3</td>
<td>sleep [n.]</td>
<td>40</td>
<td>0.56</td>
<td>3.85</td>
<td>sleeper</td>
<td>10</td>
<td>1.13</td>
<td>4.29</td>
</tr>
<tr>
<td>4</td>
<td>alarm</td>
<td>10</td>
<td>0.45</td>
<td>3.64</td>
<td>asleep</td>
<td>104</td>
<td>1.07</td>
<td>4.23</td>
</tr>
<tr>
<td>5</td>
<td>dreams</td>
<td>10</td>
<td>0.41</td>
<td>3.54</td>
<td>mornings</td>
<td>30</td>
<td>1.05</td>
<td>4.22</td>
</tr>
<tr>
<td>6</td>
<td>morning</td>
<td>80</td>
<td>0.40</td>
<td>3.52</td>
<td>slept</td>
<td>40</td>
<td>0.55</td>
<td>3.56</td>
</tr>
<tr>
<td>7</td>
<td>wake</td>
<td>10</td>
<td>0.39</td>
<td>3.50</td>
<td>nap</td>
<td>11</td>
<td>0.53</td>
<td>3.53</td>
</tr>
<tr>
<td>8</td>
<td>dream</td>
<td>10</td>
<td>0.22</td>
<td>2.94</td>
<td>morning</td>
<td>502</td>
<td>0.49</td>
<td>3.46</td>
</tr>
<tr>
<td>9</td>
<td>suddenly</td>
<td>22</td>
<td>0.20</td>
<td>3.50</td>
<td>alarm</td>
<td>28</td>
<td>0.46</td>
<td>3.39</td>
</tr>
<tr>
<td>10</td>
<td>when</td>
<td>382</td>
<td>0.18</td>
<td>2.75</td>
<td>sleep [n.]</td>
<td>130</td>
<td>0.40</td>
<td>3.25</td>
</tr>
</tbody>
</table>
The concordance lines containing asleep/sleep and [AWAKE] are not different to those of [WAKE] and those two terms. Therefore the examples below contain other words from the semantic field of ‘sleep’:

(032) I slept and woke to find the house quiet, settled from the anger that had shaken it. (BNC fic.)
(033) Hal dozed off, waking every half hour or so. (CAE fic.)
(034) Sally, a light sleeper, was woken by the sound of footsteps around the house. (CAE fic.)

Next to antonymic expressions, [WAKE] occurs quite frequently with expressions of time, forming temporal meanings like the two below:

(035) Mornings I woke at five to pray. (CAE fic.)
(036) [T]he 75year-old lady suddenly woke up, rubbed her eyes and demanded […]. (BNC newsp.)

In contrast to the findings related to [AWAKE] above, the relation between the verb wake and temporal expressions does not seem to be very close. Fewer items were found, and the MI score of most of the collocates is lower compared to those of the collocates of [AWAKE]. The noun morning for instance has an MI score of 3.52 in combination with the verb wake; with awake however, its score is slightly higher at 4.16. Nonetheless, wake occurs far more often in both corpora than awake; the infrequency of the latter is almost problematic in terms of determining behavioural characteristics. We simply cannot draw far-reaching conclusions about the collocations of [AWAKE], given the differences between their frequencies and the frequencies of the collocates of [WAKE]. It can be said however, that both verbs have a tendency to be used in temporal and antonymic expressions.

Finally, one big difference in the collocational behaviour of [AWAKE] and [WAKE] can be detected. In Table 5.8.b above, three words can be found that are actually related to the process of ‘waking up’ itself. First of all, the top ten lists of both corpora contain the noun alarm with a MI score of 3.64 and 3.39 respectively. Secondly, in the AmE material, the form stirs, 3rd person present form of the verb stir, was found. With the MI score of 4.46, it is the second of the top ten collocates of [WAKE] in the CAE, and therefore closely bound to the verb wake. Furthermore, in the BNC, the verb wake itself is listed as a collocate of [WAKE]:


It was still dark when the fire **alarm woke** him. (CAE fic.)

The guard **stirs** and almost **wakes** up. (CAE fic.)

"**Wake** up! **Wake** up!" she cried, as she rubbed his icy cheeks to warm them and try to bring him round. (BNC fic.)

We can see in (037) and (038) above that the words *alarm* and *stirs* are related to the process of waking: usually, *alarm* is related to the cause of waking up, and the verb *stir* refers to the process of waking up itself. In (039), *wake* collocates with itself for the purpose of emphasis. This is a quite frequent way to use this verb, even in the AmE variety. In the CAE, *wake*, with a MI score of 3.19, collocated with [WAKE] 54 times. This was however not enough for *wake* to occur in the top ten list.

### 5.8.2 Post-node collocates of (a)wake

Table 5.8.c below presents the post-node collocates of the verb *awake* in the BNC and the CAE. Regarding the BNC, the same problem occurs as with the pre-node collocates of [AWAKE] discussed above. Too few examples of [AWAKE] exist in the BrE material, so that personal pronouns (like *him, her, she*), prepositions (like *with, by*) and even auxiliaries (e.g. *was*) occur among the top ten collocates, although they have a rather low MI score and therefore no close connection to the verb *awake*.

<table>
<thead>
<tr>
<th>#</th>
<th>collocate</th>
<th>BNC F</th>
<th>%</th>
<th>MI</th>
<th>collocate</th>
<th>CAE F</th>
<th>%</th>
<th>MI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>morning</td>
<td>28</td>
<td>0.14</td>
<td>4.49</td>
<td>dawn</td>
<td>18</td>
<td>0.19</td>
<td>4.60</td>
</tr>
<tr>
<td>2</td>
<td>find</td>
<td>22</td>
<td>0.05</td>
<td>3.54</td>
<td>morning</td>
<td>134</td>
<td>0.13</td>
<td>4.22</td>
</tr>
<tr>
<td>3</td>
<td>early</td>
<td>11</td>
<td>0.03</td>
<td>3.05</td>
<td>a.m.</td>
<td>21</td>
<td>0.09</td>
<td>3.82</td>
</tr>
<tr>
<td>4</td>
<td>from</td>
<td>38</td>
<td>0.01</td>
<td>1.79</td>
<td>sound</td>
<td>39</td>
<td>0.07</td>
<td>3.58</td>
</tr>
<tr>
<td>5</td>
<td>him</td>
<td>10</td>
<td>0.01</td>
<td>1.43</td>
<td>find</td>
<td>91</td>
<td>0.06</td>
<td>3.42</td>
</tr>
<tr>
<td>6</td>
<td>with</td>
<td>34</td>
<td>0.01</td>
<td>1.22</td>
<td>Monday</td>
<td>11</td>
<td>0.05</td>
<td>3.33</td>
</tr>
<tr>
<td>7</td>
<td>her</td>
<td>14</td>
<td>&lt;0.01</td>
<td>1.08</td>
<td>feeling</td>
<td>21</td>
<td>0.05</td>
<td>3.17</td>
</tr>
<tr>
<td>8</td>
<td>by</td>
<td>20</td>
<td>&lt;0.01</td>
<td>0.93</td>
<td>night</td>
<td>70</td>
<td>0.04</td>
<td>3.14</td>
</tr>
<tr>
<td>9</td>
<td>she</td>
<td>11</td>
<td>&lt;0.01</td>
<td>0.69</td>
<td>start [n.]</td>
<td>36</td>
<td>0.04</td>
<td>3.09</td>
</tr>
<tr>
<td>10</td>
<td>was</td>
<td>27</td>
<td>&lt;0.01</td>
<td>0.67</td>
<td>next</td>
<td>57</td>
<td>0.03</td>
<td>2.78</td>
</tr>
</tbody>
</table>

The top four items of the BrE material however seem to be related to the verb *awake*. For instance the noun *morning* and the adverb *early* have to be recognised as collocates of [AWAKE]. They have relatively high MI scores (4.49 and 3.05 respectively), and they belong to the concept of ‘time’, which was already recognised as one of the semantic fields [AWAKE] is preferably connected with.
Even more of those terms related to the concept of ‘time’ can be found in the AmE material. The nouns *dawn* and *Monday*, as well as the abbreviation *a.m.* (meaning ‘before noon’) are combined with [AWAKE] to form temporal expressions:

(040) **We awoke Monday** morning to pouring rain. (CAE misc.)

(041) **I awake** at six *a.m.* to discover my place is in splendid disarray. (CAE fic.)

Some words can be found in Table 5.8.c above that seem to be part of fixed phrases including [AWAKE]. One of them is for instance the verb *find*, which occurs in the BNC (with a MI score of 1.79) and in the CAE (with a MI score of 3.42). In the company of [AWAKE], *find* is quite frequent (38 occurrences in the BNC and 91 in the CAE), used in a fixed phrase *[AWAKE] to find*, which means ‘to awake and realise something’:

(042) **I once met a girl who awoke to find** herself gripped by the Spirit of God at 2 a.m. (BNC misc.)

Another noun [AWAKE] combines with to form a fixed phrase is *start*. It is only present in the top ten of the CAE, having an MI score of 3.09. It is used in the phrase *[AWAKE] with a start*, meaning ‘to suddenly startle from sleep’:

(043) **John awoke with a start.** “Huh?” (CAE fic.)

Table 5.8.d below presents the top ten post-node collocates of the verb *wake* in the BNC and the CAE. Here, we can observe similar tendencies as in the discussion of the pre-node collocates of [WAKE] above.

<table>
<thead>
<tr>
<th>#</th>
<th>BNC</th>
<th></th>
<th></th>
<th>CAE</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>collocate</td>
<td>F</td>
<td>%</td>
<td>MI</td>
<td>collocate</td>
<td>F</td>
</tr>
<tr>
<td>1</td>
<td>REM</td>
<td>12</td>
<td>3.81</td>
<td>5.78</td>
<td>goggly</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>up</td>
<td>1891</td>
<td>1.05</td>
<td>4.48</td>
<td>refreshed</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>morning</td>
<td>150</td>
<td>0.75</td>
<td>4.15</td>
<td>up</td>
<td>12850</td>
</tr>
<tr>
<td>4</td>
<td>sleeping [adj.]</td>
<td>17</td>
<td>0.69</td>
<td>4.07</td>
<td>nap</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>dawn</td>
<td>12</td>
<td>0.51</td>
<td>3.77</td>
<td>disoriented</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>up to</td>
<td>118</td>
<td>0.49</td>
<td>3.73</td>
<td>nightmares</td>
<td>22</td>
</tr>
<tr>
<td>7</td>
<td>wake</td>
<td>11</td>
<td>0.43</td>
<td>3.59</td>
<td>trance</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>sleep [n.]</td>
<td>28</td>
<td>0.39</td>
<td>3.49</td>
<td>dawn</td>
<td>87</td>
</tr>
<tr>
<td>9</td>
<td>early</td>
<td>86</td>
<td>0.26</td>
<td>3.08</td>
<td>gasping</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>find</td>
<td>86</td>
<td>0.21</td>
<td>2.88</td>
<td>shivering</td>
<td>15</td>
</tr>
</tbody>
</table>
First of all, similarly to the discussion about the pre-node and post-node collocates of [AWAKE] we can see that also the verb *wake* frequently co-occurs with items belonging to the concept of ‘time’. Those are the nouns *dawn* (present in the top ten list of the BNC and the CAE) and *morning* as well as the adverb *early*. All three items were already found and discussed in relation to the verb *awake* above. Similarly to the discussion of the pre-node collocates of [WAKE] in Section 5.8.1, the verb *wake* itself was found as a collocate of [WAKE] in the BrE material. The fact that the verb *wake* occurs among the top ten collocates of [WAKE] in both pre-node and post-node positions allows the conclusion that emphasised phrases like *Wake Up! Wake Up!* (see example (039) above) are quite frequent. They could possibly be recognised as fixed phrases as well.

In contrast to the discussion of post-node collocates of [AWAKE], but similarly to the discussion of pre-node collocates of [WAKE], post-node collocates of the verb *wake* were found that belong to the semantic field of ‘sleep’. The abbreviation *REM* (meaning *rapid eye movement*) is the most strongly bound post-node collocate in the BNC, reaching a MI score of 5.78, which suggests a strong connection between this word and [WAKE]. However, all twelve instances found belong to one and the same text, and therefore one might suggest that it was the personal preference of the author rather than a specific characteristic of the abbreviation *REM* that led to the term being commonly combined with the verb *wake*. Together with the noun *sleep*, another post-node collocate present in the BrE top ten list above, it is used in rather academic phrases exemplified by (044) below. Furthermore, the AmE top ten list contains three more items related to the concept of ‘sleep’. Those are the nouns *nap*, *nightmares* and *trance*, the latter of which refers to a sleep-like status, often related to the phenomenon of hypnosis or similar. Together with items belonging to the semantic field of ‘sleep’, the verb *wake* is used to create antonymic expressions, like in the examples below:

(044) Dement’s discovery was that when subjects were *woken* during REM *sleep* they nearly always reported that they were dreaming. (BNC acad.)

(045) She began to *wake* less with *nightmares*, at least the screaming kind. (CAE fic.)

(046) At the last possible moment, Nick seems to *wake* from his *trance*. (CAE fic.)
Combined with [WAKE], the verb *find* (present in the BNC top ten list with a MI score of 2.88) forms similar expressions as it does with [AWAKE] (discussed above). Together with [WAKE] and the preposition *up* (present in both corpora in Table 5.8.d), a very frequent collocate of the verb *wake, find* forms the fixed expression *[WAKE] up to find*, a phrase that is synonymous to *[AWAKE] to find*, bearing the meaning ‘to awake and realise something’:

(047)  Charsky woke up to find his father, vividly present a moment before, vanished from the room. (BNC fic.)

Finally, a big difference to the post-node collocates of [AWAKE] was found. In the AmE material, five different adjectives can be found that tend to co-occur with [WAKE] in post-node positions: *groggy, refreshed, disoriented, gasping, and shivering*. Similarly to the observations made for the verb *rise* in Section 5.7.2, it can be concluded that *wake* is often used in descriptive sentences, though in a slightly different way than the two verbs discussed above:

(048)  Nutmeg is a strong sedative, so it's possible to wake up groggy from too much. (CAE misc.)
(049)  Tim feels that he's waking up more refreshed, he feels better […]. (CAE spok.)
(050)  Chad woke up disoriented, with a headache like a construction crew working between his ears. (CAE fic.)
(051)  Every time I'd doze off, I'd wake up gasping for air, and I said anything's better than that. (CAE spok.)

In (048) – (051) above we can see that *wake* in combination with those adjectives describes the state in which a person was after waking up, not the process of waking up itself. Most of the phrases relate to the feelings of the subjects, but those feelings are closely bound to the verb: *wake up groggy, waking up […] refreshed, woke up disoriented, wake up gasping*. Albeit this difference to the descriptive use of the verbs *arise and rise*, [WAKE] is used to form descriptive sentences as well.

5.9 Summary: the collocates of (a)rise and (a)wake

In Section 5.7 and 5.8, the top ten collocates of the verbs *arise, rise, awake and wake* have been investigated. The results of the discussions are summarised in this section.

In Section 5.7 it was shown that the verb *arise* very often combines with negative words in pre-node position. [ARISE] itself might therefore have a specific semantic prosody, being automatically associated with negative feelings. In post-node position, [ARISE] tends to
combine with prepositions or prepositional phrases. In those cases, the verb *arise* is part of causative expressions. Furthermore, [ARISE] frequently combines with adverbs, creating descriptive senses. This observation is probably applicable to most verbs, however. Some frequent fixed phrases containing *arise* were also found, e.g. [ARISE] *from the ashes*, which is a synonymous variant of [RISE] *from the ashes*.

The verb *rise* also very frequently co-occurs with nouns in pre-node position. Those nouns do not have a specific semantic prosody like the collocates of *arise*. However, most of the pre-node collocates of [RISE] belong to the semantic field of ‘economy’. As post-node collocates, the verb *rise* prefers adverbs, creating expressions of manner like it has been observed for [ARISE] as well. Some of the fixed phrases that could be identified for the verb *rise* are [RISE] *to a crescendo* and [RISE] *to prominence*.

Finding out about the collocates of [AWAKE] proved to be a bit difficult, due to the scarcity of occurrences, especially in the BNC. It was found, however, that the verb *awake* prefers words related to the concepts of ‘sleep’ in pre-node, and to the semantic field of ‘time’ in pre- and post-node positions. Evidence for some fixed phrases like [AWAKE] *to find* (a synonym to [WAKE] *up to find*) and [AWAKE] *with a start* were also found.

Similarly to the verb *awake*, [WAKE] prefers words belonging to the semantic field of ‘sleep’ and ‘time’ in pre-node and post-node position. Furthermore it occurred in pre-node and post-node position that [WAKE] collocates with itself to form emphasised phrases like *Wake up! Wake up!* In post-node position, especially in the AmE material, [WAKE] frequently co-occurs with adverbs and adjectives, creating descriptive senses similar to the verbs *arise* and *rise*.

### 6. Qualitative analysis

This section contains the semantic analysis of 250 concordance lines of each verb. Before presenting and discussing the results, some words have to be said about the selection method of the material that was investigated. The following Section 6.1 deals with the investigation of the meanings of *arise* and *rise*, whereas Section 6.2 presents the results for *awake* and *wake*. Due to the lack of space and time, it was decided to investigate not more than 250 concordance lines of each verb. The selection of those concordance lines was governed by some conditions. It was decided to select the material according to the proportional size of the different subcorpora of the BNC and the CAE.
The different subcorpora of the BNC and the CAE are quite identical in their characteristics, for instance in the type of texts they contain. Therefore, it was decided to use the term *genre* instead of *subcorpus*, since the concept of *genre* can be easily related to semantic analyses like this one. The percentage column is most important in Table 6.a. Here we can see the proportion according to which the material for this analysis has been selected. This column shows that the language genres are quite evenly distributed within the two corpora. Most of the genres contain around 19% of the whole 467.3 million words. Only the miscellaneous category is a bit overrepresented, containing 26.2% of all words stored in the BNC and the CAE. For the process of selecting material for the present qualitative analysis, these percentages mean that for each verb, 47 (18.6%) of the 250 concordance lines were taken from the spoken language subcorpus of the BNC and the CAE, 46 instances (18.3%) were taken from the fiction category of both corpora, etc.

Secondly, a differentiation was made between the BrE and the AmE material. For the purpose of accuracy, the sizes of the subcorpora of the BNC and the CAE have been put into relation to the whole of the data used in this study. Table 6.b below presents the actual numbers of concordance lines, their distribution within the different corpora and their subcategories. To understand the content of Table 6.b one has to consult Table 6.a as well. There we can see that the total 86.9 million words within the spoken language genre are unevenly distributed between the BNC and the CAE. The 10.3 million words stored in the spoken language subcorpus of the BNC result in 11.85% of the whole 86.9 million words in the spoken language genre. Those 11.85% have also been taken into account in the process of material selection. It means that 11.85% of the 47 concordance lines belonging to the spoken language genre were taken from the spoken language subcorpus of the BNC. The remaining 88.15% of spoken language concordance lines were then taken from the CAE spoken language subcorpus. The other proportions that were taken into account during the selection of the material can be seen in Table 6.b. The uneven distribution of

<table>
<thead>
<tr>
<th>Genre</th>
<th>Size in MW</th>
<th>BNC</th>
<th>CAE</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td>10.3</td>
<td>76.6</td>
<td>86.9</td>
<td></td>
<td>18.6</td>
</tr>
<tr>
<td>Fiction</td>
<td>15.9</td>
<td>69.6</td>
<td>85.5</td>
<td></td>
<td>18.3</td>
</tr>
<tr>
<td>Newspaper</td>
<td>10.5</td>
<td>73.4</td>
<td>83.9</td>
<td></td>
<td>18.0</td>
</tr>
<tr>
<td>Academic</td>
<td>15.3</td>
<td>73.0</td>
<td>88.3</td>
<td></td>
<td>18.9</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>44.6</td>
<td>78.1</td>
<td>122.7</td>
<td></td>
<td>26.2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>96.6</strong></td>
<td><strong>370.7</strong></td>
<td><strong>467.3</strong></td>
<td></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genre</th>
<th>N</th>
<th>BNC</th>
<th>%</th>
<th>CAE</th>
<th>%</th>
<th>TOTAL</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td>6</td>
<td>11.85</td>
<td></td>
<td>41</td>
<td>88.15</td>
<td>47</td>
<td>18.8</td>
<td></td>
</tr>
<tr>
<td>Fiction</td>
<td>9</td>
<td>18.60</td>
<td></td>
<td>37</td>
<td>81.4</td>
<td>46</td>
<td>18.4</td>
<td></td>
</tr>
<tr>
<td>Newspaper</td>
<td>6</td>
<td>12.50</td>
<td></td>
<td>39</td>
<td>87.5</td>
<td>45</td>
<td>18.0</td>
<td></td>
</tr>
<tr>
<td>Academic</td>
<td>8</td>
<td>17.30</td>
<td></td>
<td>39</td>
<td>82.7</td>
<td>47</td>
<td>18.8</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>24</td>
<td>36.35</td>
<td></td>
<td>41</td>
<td>63.65</td>
<td>65</td>
<td>26.0</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>53</strong></td>
<td><strong>197</strong></td>
<td></td>
<td><strong>250</strong></td>
<td></td>
<td><strong>100.0</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
BNC and CAE material leads to an overrepresentation of AmE, 79% of the concordance lines investigated belong to this language variety. However, this method of material selection can still be regarded as suitable for the present investigation, first of all because the differences between BrE and AmE cannot be investigated by analysing 250 instances per verb, even if those were evenly distributed. Larger investigations would be necessary to compare the semantic values of the verbs in both language varieties. Hence, this qualitative analysis will not focus on the differences between BrE and AmE, but on the meaning of the verbs in question within different language genres. Furthermore, the proportional selection of BrE and AmE material reflects the difference in size that exists between both corpora used in this study. The third and last condition that governed the material selection for this analysis is that all concordance lines were selected randomly. It was made sure that no personal preferences interfered during the process of selecting material.

The tables in Section 6.1 and 6.2 refer to the meanings shared by **arise** and **rise** and **awake** and **wake** respectively. The bases especially of Table 6.1.a and 6.2.a are the meaning descriptions for each verb in the OED. The meanings of **arise** and **rise** and **awake** and **wake** respectively have been compared with each other and similar meanings were grouped together. Hence, it was found that **arise** and **rise** share 18 meanings, while **awake** and **wake** share six.

**6.1 Meanings of (a)rise**

Table 6.1.a presents the meanings shared by **arise** and **rise** and their distribution among 250 examples of each verb. Together with Table 6.1.b an attempt is made to characterise and exemplify the shared meanings of **arise** and **rise**. Table 6.1.a as well as Table 6.2.a in the next section contain a column that needs to be explained briefly. In both tables information about tags is given, referring to the current status of the use of the verb in question in a specific meaning. Those tags were taken from the OED and the abbreviations listed there mean the following: A = archaic, O = obsolete, P = poetic and R = rare.

Table 6.1.b below shows the frequency of the meanings of **arise** and **rise** within the different language genres. Comments on this table will be made whenever it is necessary to refer to the distribution of a meaning category among the different language genres. Below a more general overview of the meanings of **arise** and **rise** is given, followed by a more detailed discussion of the most frequent meanings of both verbs in Section 6.1.1.
It can be seen in Table 6.1.a above that the majority of meanings are more frequently expressed with the verb *rise*. Seven meanings are more frequently expressed with *arise* instead, although one has to be careful with general conclusions since many of those meanings did not occur very frequently (meaning 05 for instance has only three instances of [ARISE] and one of [RISE]).
### Table 6.1.b. Distribution of the meanings of [ARISE] and [RISE] within the different language genres

<table>
<thead>
<tr>
<th>#</th>
<th>Genre</th>
<th>ARISE</th>
<th></th>
<th></th>
<th></th>
<th>RISE</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>02</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>03</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>05</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>06</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>07</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>08</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>09</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>29</td>
<td>0</td>
<td>19</td>
<td>10</td>
<td>15</td>
<td>73</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>9</td>
<td>19</td>
<td>13</td>
<td>47</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>6</td>
<td>8</td>
<td>11</td>
<td>13</td>
<td>43</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>18</td>
<td>35</td>
<td>26</td>
<td>30</td>
<td>27</td>
<td>43</td>
<td>161</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>46</td>
<td>45</td>
<td>47</td>
<td>65</td>
<td>250</td>
<td>47</td>
<td>46</td>
<td>45</td>
<td>47</td>
<td>65</td>
<td>250</td>
</tr>
</tbody>
</table>

With regard to two meanings it can be said that [ARISE] and [RISE] occur equally often. First of all, meaning 02, ‘to get up from a fall’, was not present among the 250 concordances lines of either verb. This meaning is tagged as archaic for arise, but it seems to be an uncommon sense of rise as well. The second meaning that is almost equally common for arise and rise is number 03, ‘to get up from sleep or rest’. Seven occurrences with [ARISE] and nine with [RISE] were found. Although, similarly to meaning 02, this meaning of arise has been tagged as archaic, the verb seems to be quite frequently used in this sense. Another similarity between arise and rise expressing meaning 03 can be seen in Table 6.1.b. The majority of instances for both verbs were found in texts belonging to the fiction genre. The two examples below show that no difference in meaning is noticeable, and that both verbs can be used interchangeably in this sense. It can therefore be concluded that the verbs arise and rise are synonyms when expressing meaning 03:

(052)  “Nap? But you said he had just arisen”. “He liked a nap soon upon awakening, mistress”. (CAE fic.)

(053)  He slept fitfully, rising in the cold rain of morning to see how much the desert had weakened around [...]. (CAE news.)
The next group of senses discussed here include meaning 05, 08 and 15. All three favour the use of *arise*, but they are quite infrequent in general. Instances of group 05 express the meanings ‘to rise from the dead’ and ‘return from the grave’, and more instances with *arise* (i.e. 3 against 1 *rise*) were found. As we can see in Table 6.1.a above, this meaning is tagged as poetic for *arise*, and accordingly, the three instances were found in the fiction and spoken language genres, both of which do indeed contain poetic language. Examples (054) and (055) describe the use of *arise* and *rise* expressing meaning 05:

(054) In church we sang “Up from the grave he *arose*”, and my mother and brother started laughing. “He *arose!* He *arose!*”. (CAE fic.)
(055) [B]asketball has *risen* from its death bed to become the world's hottest sports business […]. (CAE news.)

Meaning 08 refers to sounds (‘to come up aloud; to be heard; to get louder’) and, having the same frequencies as meaning 05, it is as well more commonly expressed with *arise*. Interestingly, this meaning 08 of *arise* is also tagged as being archaic, although it is so frequent that it does occur among the small sample of 250 concordance lines. (056) and (057) below exemplify that this meaning can be expressed in two different ways. First of all, it can be directly related to the growing of sounds, secondly it can refer to specific feelings that more or less directly are related to sounds:

(056) A sympathetic murmur *arose* from the crowd. (CAE fic.)
(057) Public concern was growing; in some quarters, panic and hysteria *arose*, or were fanned. (CAE misc.)

Meaning 15 fits into this part of the discussion as well. Only two instances were found, both contain the verb *arise* expressing the sense ‘to be born; come into the world of life or action’ in reference to persons or characters. Although this meaning seems to be very infrequent it has not been tagged as archaic by the OED. Example (058) displays the use of *arise* in this sense, although it does not refer to a single person or character but to a species in general:

(058) […] but as one species died out, a very similar one would *arise* to take its place. (CAE misc.)

Meaning 06, in contrast to the other meanings already investigated, is preferably expressed with the verb *rise*, nine instances of which could be found. Only three occurrences with
[ARISE] were counted in the 250 concordance line sample. The sense of this group is ‘to rise from inaction; to rise up against; to rebel’, and three different ways of expressing those meanings were found. (059) below displays the most frequent usage of meaning 06, expressing ‘to offer resistance’, often used in imperative phrases. (060) then refers to the ‘rise from inaction’, of which only one example could be found. Finally, (061) displays the sense ‘to take hostile steps against someone or something’.

(059) Women opera-goers arise! (CAE news.)
(060) […] persuaded ourselves were eager to save us the trouble of invading, didn't rise up to help our proteges at the Bay of Pigs […]. (CAE misc.)
(061) In some countries rebels make up their own military force and rise up against the government. This is usually triggered by disagreements over their rights. (BNC misc.)

The last category that will be discussed in this section will be that of ‘other meanings’, containing senses that are not shared by arise and rise. Only instances of rise could be found which belong to this category, since they express meanings that cannot be expressed with the verb arise as well. Table 6.1.a above shows that fourteen instances were put into the other category, resulting in almost six percent of all 250 concordance lines – an observation which allows the conclusion that the semantic range of rise is bigger than that of arise. Those fourteen instances are distributed quite evenly within four different types of meanings in which only the verb rise can be used. (062) – (065) below exemplify those four different meanings. (062) shows that rise can be used as a synonym for ‘jump higher than someone or something’, an expression that probably originates in the sports genre. (063) refers to the verb rise in the sense of ‘to accept, deal with something’, while (064) is an expression related to the process of ‘spotting or seeing something’. Finally, (065) displays the use of rise in the fixed phrase [RISE] to the occasion:

(062) Liam Herbert, one of the shortest players on the pitch, rose above the defenders and placed a fine header past keeper Duncan Green […]. (BNC spok.)
(063) West Indian cricket has grown strong precisely because it has risen above its problems […]. (BNC misc.)
(064) As I reach this place on the trail, my eye rises above a feathery late-summer tangle of wild raspberry, cow parsnip, and mountain hollyhock […]. (CAE fic.)
(065) TRYTON INN rises to the occasion! Now you can "Snack" without that guilty feeling […]. (BNC misc.)
Even though some dissimilarities between the meanings of the two verbs could be found, in general one can say that \textit{arise} and \textit{rise} complement each other; meanings where \textit{arise} is infrequent show high frequency of \textit{rise} instead, and the same holds for the other way around. Some meanings in Table 6.1.a above can be characterised as “typical” for either \textit{arise} or \textit{rise}. Those meanings are discussed in Section 6.1.1.

\textbf{6.1.1 The most frequent meanings of (a)rise}

The meanings at the bottom of Table 6.1.a are dominated by the use of \textit{arise}. The four meanings 14, 16, 17 and 18 are the most frequent and most typical senses [ARISE] can express, and they are discussed in more detail below.

All four meanings that are most frequently expressed with \textit{arise} relate to the origin of something. They differ basically in their relation to different concepts. Meaning 14 for instance relates to the origin of something from a specific source, like rivers which spring forth from their sources. Seven instances with [ARISE] were found, while no instances of this meaning with [RISE] occurred among 250 concordance lines. Therefore, it is probably one of the major meanings \textit{arise} is used in. The majority of instances (as can be seen in Table 6.1.b) were found in the miscellaneous subcorpora and in the newspaper genre. Examples (066) and (067) below display the two different ways \textit{arise} can be used to express meaning 14. First of all, in a more literal sense, a reference to a location can be expressed. Furthermore, meaning 14 is more frequently used in a figurative way in order to refer to an abstract concept as source:

\begin{itemize}
\item (066) If a precise diagnosis is made and the problem is in fact \textbf{arising from} the joint, not the muscles where it is most common, then surgery […]. (CAE misc.)
\item (067) Peru: the very name \textbf{arose from} a misunderstanding on the part of the Spaniards when Indians told them the name of a river […]. (BNC misc.)
\end{itemize}

Meaning 16, ‘to spring up; to be raised or build’, is very similar to meaning 14 and even 15 (discussed above), with the difference that the focus is not on persons or characters but on concrete or abstract concepts. Only one instance with [RISE] was found. It seems as if this meaning is another typical sense expressed with the verb \textit{arise} instead: nineteen instances were found, the majority of which occurred in the academic language genre. This is quite surprising since this meaning of \textit{arise} is tagged as poetic by the OED. (068) and (069) below exemplify the use of \textit{arise} and \textit{rise} in this sense:
As his series of ten or so articles proves, the masculine and neuter *arose* before the feminine; the latter is a third gender, an historical afterthought. (CAE fic.)

Major stages characterize the rise and fall of empires throughout history: an empire *rises* as a result of a vacuum left by declining empires […]. (CAE misc.)

The two meanings most frequently expressed with [ARISE] are meaning 17 and 18. Table 6.1.a shows that 17.2 percent of all concordance lines belong to the former and almost 65% to the latter group. The two meanings are quite similar, and they are also related to the sense of meaning 16. Meaning 17, ‘to spring, originate or result from something’, focuses on the origin of something as a result of something else, whereas meaning 18, ‘to present itself; to come into existence’, focuses on the coming into existence itself. Meaning 17, due to the high representation, is one of the typical senses that are expressed with *arise* while only one instance with [RISE] was found. Instances of [ARISE] expressing meaning 17 are present in all language genres, which is another indication of the rather general applicability of this meaning of the verb *arise*. With [ARISE], this meaning refers almost exclusively to abstract concepts, as (070) below shows. It also exemplifies the causative usage of *arise* in combination with one of the most frequent pre-node collocates of the verb *arise*, discussed in Section 5.7.1:

> Miss T.’s father was anxious as to any complications which might *arise from* her mother’s religious beliefs […]. (BNC acad.)

Meaning 18 is even more frequent than meaning 17: 161 instances with [ARISE] were found. They are quite evenly distributed among the different language genres, which allows the conclusion that meaning 18 is the most typical meaning of [ARISE]. Meaning 18 is related to rather abstract things or concepts. Nouns like *question, issues* or other words with negative connotations often collocate with *arise* in this sense (see Section 5.7 for more information), as is exemplified below:

> […] with this kind of behaviour, a collective anti-German consciousness *arose* I think which, strangely, even today unites Russians, Poles […]. (BNC spok.)

The verb *rise* is not very common in this sense. Only four of the 250 concordance lines belonged to this group; two of them are presented below. There it can be noticed that, in
contrast to *arise, rise* in this sense is related to rather concrete concepts (things you can recognise or feel), like *tears* and *feeling*:

(072) As she watched them, Helen felt tears *rise* to her eyes, and she had no idea why. (CAE fic.)
(073) Among us Europeans who were traveling together in America […] there *rose up* repeatedly this pharisaical feeling: we all have something that you lack […]. (CAE acad.)

Next the most frequent meanings expressed with *rise* are discussed. The first meaning displayed in Table 6.1.a above is one of those examples where *rise* is quite often used. The verb *arise* however is tagged as archaic in this sense, hence it is not surprising that only 2 instances could be found. Meaning 01 refers to the process of ‘getting up from sitting or kneeling; to get upon one’s feet’, and the majority of instances were found in the fiction genre. Furthermore, this meaning refers to two different but related senses. First of all, it most frequently refers to persons or characters that get up from sitting (see the first example below). Secondly, and very rarely, it is used in reference to animals that “assume an erect position on the hind legs” (OED), exemplified by (075) below (and example (018) as well):

(074) Rachel finally *rises* from the sofa. (CAE misc.)
(075) The bear comes into full view […]. Instantly he stops, then *rises* on his hind legs […]. (CAE fic.)

Meaning 04 occurred twelve times with the verb *rise*, while no instance with [ARISE] was found. This was not unexpected, since the OED tagged *arise* as archaic and poetic in this sense. The meaning ‘to come above the horizon’ was found most often within the newspaper and academic language genre (as can be seen in Table 6.1.b). Two different senses can be expressed. On the one hand, a direct reference to the rise of heavenly bodies, mostly the sun, can be expressed, like in (076). On the other hand, there can be a more figurative sense, with the meaning ‘to spread out over the horizon’, as the darkness in example (077) below:

(076) Because the sun has *risen* in the sky, I roll my shirtsleeves up. (CAE fic.)
(077) The black horizon began to *rise*, to mount into black ridges and hills. (CAE fic.)

Meaning 07, ‘to rise in violence or agitation; to increase in force’, is tagged as poetic for *arise* and was not found among 250 concordance lines. The eight instances with [RISE] are evenly distributed among the different language genres; therefore it seems to be a rather typical sense.
of this verb. Similarly to many of the meanings already discussed, two different ways of expressing this meaning are possible. Literally, it can be related to the upcoming or intensification of wind, as in example (078) below. Figuratively, it relates to things or often feelings that become more intense or strong in general, like the heat in (079):

(078) Outside, the wind has **risen**. The forecaster promised a calm day. (CAE fic.)
(079) Fierce heat began in spring and kept **rising**.

Meaning 09, ‘to go up, ascend into the air’, occurred four times among the 250 concordance lines with [ARISE]. Suitably, the majority of those instances occurred in the fiction genre, since meaning 09 is characterised as poetic by the OED. Even more frequently this meaning is expressed with the verb **rise**; sixteen instances (6.4 %) were counted. Similarly to meaning 07, meaning 09 can be expressed in a literal and a figurative way, the former relating to smoke or dust ascending into the air (080), the latter expressing something coming up to a surface:

(080) Till dust **arises**, and grass laid flat […]. (CAE fic.)
(081) I placed a foot on a boulder that had **risen** like an idol's skull in the runoff of a heavy rain. (CAE fic.)

Meaning 10 (‘to rise with its summit or surface; to swell; to grow in respect of height’), is one of the groups that are tagged as obsolete senses for the verb **arise**. Accordingly, no instances were found among 250 concordance lines. With [RISE], only seven instances were counted, most of which, according to Table 6.1.b, occurred in the fiction corpus. Once again, two ways of expressing this meaning were found. First of all, as can be seen in (082) below, a reference to water or the water-level can be made. Secondly, it can refer to things that ‘puff out’ due to the activity of yeast or similar substances, like the expression in (083) shows:

(082) From the point when the water first burst the banks and began to **rise up** the gardens of the properties it was mainly a battle against he [sic] clock […]. (BNC misc.)
(083) […] but you forgot to mix lard in, and the salt and baking powder, and the biscuits didn't **rise** […]. (CAE fic.)
The remaining three meanings 11, 12 and 13 share some characteristics. Not only are they all related to the concept of ‘increasing value’, but they are also tagged as obsolete senses for the verb *arise*. It is therefore somewhat surprising that one instance each of meaning 12 and 13 containing this verb were found.

Meaning 11 expresses ‘to rise in rank or eminence’ and it is almost exclusively used in reference to persons or characters. Twenty-one instances with [RISE] were found which are quite evenly distributed among the different language genres. Although the resulting percentage of 8.4 makes it “only” the fourth most frequent meaning expressed with the verb *rise*, due to the circumstances that no instances with [ARISE] occurred and the fact that the verb *arise* is tagged as obsolete in this sense, it can be said that meaning 11 is one of the major senses of the verb *rise*. Once again, two different ways were encountered to express this sense. Most frequently, like example (084) shows, the level to which someone rises is in the focus of the expression. Furthermore, meaning 11 can also be expressed by focusing on the process of becoming more important in general, like the second example below shows:

(084) Tyson, the man who **rose from** the ghettos of New York to the pinnacle of boxing success […]. (BNC news.)

Meanings 12 and 13 are very similar. Both have the sense ‘to rise or increase in price, amount or importance’ and both are very frequently expressed by means of [RISE]. Furthermore, both meanings are tagged as obsolete for *arise*, but still one instance for each meaning could be found (see (086) below). However, the difference between the two meanings is their focus. Meaning 12, ‘to rise, increase in price, amount or importance’, is very similar to meaning 11, but it basically refers to abstract concepts, not persons or characters. For the verb *rise*, 73 instances (29.2 %) were found that belong into this group, the majority of which were found in the spoken language genre (29) followed by the newspaper genre (19). Meaning 12 is the most frequent sense of the verb *rise*. It is also the meaning that most often is expressed with *rise* in the company of the most frequent collocates, e.g. *prices, incomes* and other terms belonging to the semantic field of ‘economy’ (see Section 5.7), like the noun *costs* in (087):

(086) This new community, this new kind of church, is **arising** across America, but it is clearly visible in San Francisco […]. (CAE news.)

Meanings 12 and 13 are very similar. Both have the sense ‘to rise or increase in price, amount or importance’ and both are very frequently expressed by means of [RISE]. Furthermore, both meanings are tagged as obsolete for *arise*, but still one instance for each meaning could be found (see (086) below). However, the difference between the two meanings is their focus. Meaning 12, ‘to rise, increase in price, amount or importance’, is very similar to meaning 11, but it basically refers to abstract concepts, not persons or characters. For the verb *rise*, 73 instances (29.2 %) were found that belong into this group, the majority of which were found in the spoken language genre (29) followed by the newspaper genre (19). Meaning 12 is the most frequent sense of the verb *rise*. It is also the meaning that most often is expressed with *rise* in the company of the most frequent collocates, e.g. *prices, incomes* and other terms belonging to the semantic field of ‘economy’ (see Section 5.7), like the noun *costs* in (087):

(087) Meanwhile, the cost of Medigap insurance has **risen** between 20 and nearly 70 percent this year. (CAE spok.)
Meaning 13 on the other hand relates to a specific point on a scale that something rises to. It also refers to abstract things like *numbers* or *prices*, which have been identified as frequent collocates of *rise*. It is the second most frequent sense of *rise*, represented 47 times in 250 concordance lines (18.8%). Most occurrences were found in the academic and miscellaneous language genres, resulting in a noticeable difference compared to the behaviour of *rise* in meaning 12: The two meanings seem to be prominent in different genres. Example (088) below exemplifies the most frequent and typical use of *rise* in this sense:

(088) Central shares *rose* 80p to £13.45 on the news, but most others fell. (BNC news.)

Before the meanings of the other verb pair *(a)wake* is discussed, a few words will be said about the transitivity of *arise* and *rise*.

6.1.2 *Transitivity of (a)rise*

The verb *arise* cannot be used in a transitive way. It is not possible to ‘*arise someone or something*’. Accordingly, no such instances were found among the 250 concordance lines. Regarding the verb *rise*, things are different. The OED gives the information that the transitive forms of *rise* are *raise* and *rear*, but that ‘*rise itself has to a certain extent (esp. in later use) assumed the functions of a transitive verb*’. Unfortunately, no instance of this interesting phenomenon occurred during the investigation of 250 concordance lines.

6.2 Meanings of *(a)wake*

This section discusses the distribution and characteristics of the meanings shared by *awake* and *wake*. Similarly to Section 6.1, two tables are used to summarise the findings. Table 6.2.a below presents the meanings shared by *awake* and *wake* and their distribution among 250 examples each. The following Table 6.2.b displays the frequency of the meanings of *awake* and *wake* within the different language genres. After a discussion about the rather infrequent meanings of both verbs, a more detailed discussion about their most frequent meanings will follow in Section 6.2.1.

We can see in Table 6.2.a below that, in contrast to the distribution of the shared meanings of *arise* and *rise*, the verbs *awake* and *wake* rather support each other: almost every meaning can be expressed with either [AWAKE] or [WAKE].
Table 6.2.a. Meanings shared by *awake* and *wake* and their distribution amongst 250 examples each

<table>
<thead>
<tr>
<th>#</th>
<th>OED Meaning #</th>
<th>Description</th>
<th>Tag</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>awake – 1</td>
<td>To come out of the state of sleep; to cease to sleep</td>
<td>F</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>wake – 7a,b</td>
<td>To come out of the state of sleep; to cease to sleep</td>
<td></td>
<td>130</td>
</tr>
<tr>
<td>02</td>
<td>awake – 2</td>
<td>To rise from a state resembling sleep; to become active</td>
<td></td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>wake – 7c</td>
<td>To become animate, alert or lively</td>
<td>F</td>
<td>38</td>
</tr>
<tr>
<td>03</td>
<td>awake – 3</td>
<td>To awake to sth.; to become fully conscious of</td>
<td></td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>wake – 7d,9b</td>
<td>To wake (up) to; to become conscious, aware of</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>04</td>
<td>awake – 4</td>
<td>To be or keep awake; to be vigilant; to watch</td>
<td>R</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>wake – 1,2,4,5,6,8</td>
<td>To be awake; to remain awake for a purpose; to watch</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>05</td>
<td>awake – 5</td>
<td>To arouse someone from sleep</td>
<td></td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>wake – 8</td>
<td>To rouse someone from sleep or unconsciousness</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>06</td>
<td>awake – 6</td>
<td>To rouse from a state resembling sleep; to make active</td>
<td></td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>wake – 9a,10</td>
<td>To rouse to action, activity, alertness</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other meanings</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>250</strong></td>
</tr>
<tr>
<td></td>
<td>awake</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>wake</td>
<td></td>
<td></td>
<td><strong>TOTAL</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>250</strong></td>
</tr>
</tbody>
</table>

This is not very surprising however, since this meaning has been tagged rare for the verb *awake* and obsolete for *wake* respectively. The second group that is not represented by both verbs is containing other meanings.

Table 6.2.b. Distribution of the meanings of [AWAKE] and [WAKE] within the different language genres.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>01</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>17</td>
<td>36</td>
</tr>
<tr>
<td>W</td>
<td>02</td>
<td>9</td>
<td>10</td>
<td>5</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>A</td>
<td>03</td>
<td>9</td>
<td>2</td>
<td>11</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>W</td>
<td>04</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>05</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>K</td>
<td>06</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>Other</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>47</td>
<td>46</td>
<td>45</td>
<td>47</td>
<td>65</td>
<td>250</td>
</tr>
</tbody>
</table>

As can be seen in Table 6.2.b on the left, the meaning of one concordance line from the academic language genre could not be put into one of the six categories of shared meanings of *awake* and *wake*.

Instead, this example uses an expression which – if at all – is solely bound to the verb *wake*: ‘to light (or wake) a fire’. It has to be regarded a poetic expression, see example (089) below:

(089) From the ashes a **fire shall be woken** a light from the shadows shall spring renewed shall be blade that was broken the [...] (CAE acad.)

This sense could not be found in the OED. On Google, however, some thousand instances of *wake a/the fire* can be found, many (if not most) of which are used in religious or metaphorical contexts.
6.2.1 The most frequent meanings of (a)wake

In Table 6.2.a we can see that meaning 03 is relatively frequently expressed by means of both *awake* and *wake*. [AWAKE] occurred 34 times expressing this sense (13.6 %), whereas 25 instances of [WAKE] were found (10 %). Meaning 03 has the sense ‘to awake or wake up to something; to become aware of’, and the two verbs are synonyms when expressing this meaning. Fixed phrases, like *[AWAKE] to find* (see Section 5.8.2) were frequently found among the instances belonging to meaning 03. See the examples below:

(090)  [H]e didn't like to sleep on the street because workers who do often **awake to find** their shoes, jacket or money stolen. (CAE news.)

(091)  The brewers have at last **woken up to** the fact that their high-street shops have become dinosaurs slouching towards extinction. (BNC news.)

(092)  […] so that by the time people have “**woken up** to what's happened, it's gone farther than what they feel […]”. (CAE spok.)

Examples (091) and (092) are both variants of the fixed phrase that can be found in (090). Instead, one could as well say the brewers have […] *woken up to find that* […] and people have woken up to find that it’s gone farther […], since all those expressions relate to the process of realising something.

The second most frequent meaning for *awake* and *wake* is meaning 02, ‘to rise from a state resembling sleep or inaction’, which is very frequently used in a figurative sense. Similarly to meaning 03, meaning 02 is slightly more often expressed by means of *arise*; 61 instances were found containing this verb (resulting in 24.4%), whereas only 38 instances with [WAKE] were found (15.2%). (093) – (096) below exemplify the broad range of senses that can be expressed with either verb. (093) and (094) show the most frequent figurative use of meaning 02, while (095) and (096) are rather special expressions:

(093)  But as they approached Takow, the harbor **awoke**. I vividly remember when we made our turn, I was looking ahead […]. (CAE spok.)

(094)  Morning in this street **awakes** unwashed; a stale wind breathing litter, last night’s godlessness. (BNC fic.)

(095)  But Buford finally **awoke** - and a game that should have been a thriller finally became one. (CAE news.)

(096)  [T]his is how it is, and we are to **wake up** and know it, realize it. (CAE acad.)
Example (095) represents the frequent use of [AWAKE] and [WAKE] in texts related to sports, often referring to a team that after a period of difficulties starts to play better. The following example (096) then is very similar to expressions of meaning 03: ‘to wake up and realise’.

Table 6.2.a shows that meaning 01, ‘to come out of the state of sleep’, is the most frequent meaning of both awake and wake. We can see that both verbs have a very similar frequency; around 50 percent of each sample investigated belonged to this category. Additionally, for both verbs, Table 6.2.b displays a rather even distribution of meaning 01 among the different language genres. It can therefore be concluded that in general, awake and wake most typically express the sense of meaning 01. Hence, they can be regarded synonymous in this sense. When expressing this meaning, awake and wake frequently co-occur with words referring to daytime and time in general, which were already identified as frequent pre-node and post-node collocates of both verbs (see Section 5.8). The three examples below will exemplify the use of awake and wake in this sense. (097) contains the fixed phrase [AWAKE] with a start (see Section 5.8.2), while (098) contains an antonymic expression involving the words wake and dead.

(097) When a man of some 300 pounds -- to put it conservatively -- awakes with a start, it is apt to jar the strongest chair […]. (CAE acad.)

(098) “Or wake up one morning and be dead”, added Dawn Bazner. (CAE news.)

(099) No policymaker wants to wake one day and find that the actual costs of the program are twice those originally […]. (CAE acad.)

Note that the examples (098) and (099) can easily be paraphrased by expressions fitting into meaning 03: wake up to realise that one is dead and wake to find. It seems as if instances of meaning 03 are frequently expressed by means of meaning 02 and 01 as well. This might explain why meaning 03 is rather infrequent compared to the other two groups.

6.2.2 Transitivity of (a)wake

In contrast to the verbs arise and rise, a difference between the transitive and intransitive use of the verbs awake and wake was detected. Table 6.2.c below briefly summarises the distribution of transitive and intransitive uses of [AWAKE] and [WAKE].

<table>
<thead>
<tr>
<th></th>
<th>transitive</th>
<th>intransitive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>awake</td>
<td>23</td>
<td>227</td>
<td>250</td>
</tr>
<tr>
<td>wake</td>
<td>60</td>
<td>190</td>
<td>250</td>
</tr>
</tbody>
</table>

We can see that the verb wake is more frequently used in a transitive way than awake, but that both verbs are predominantly used to express intransitive meanings.
The remaining two groups, meaning 05 and 06, will be discussed in this section, as they basically are transitive expressions of meaning 01 and 02, which only refer to the intransitive use of both verbs.

Meaning 05, ‘to arouse someone from sleep’, is more frequently expressed by means of the verb *wake*. Table 6.2.1 shows that 41 instances were found (16.4 %). Only nine instances with *awake* have been found. Hence it is more common to say ‘to wake someone up’ than ‘to awake someone’, the latter sounding a bit more poetic. In fact, as can be seen in Table 6.2.b, the majority of instances containing [AWAKE] were found in the spoken language and fiction genre, whereas for [WAKE] the instances are more evenly distributed among the different language genres. (100) – (103) below display the use of the two verbs in this sense:

(100)  When our exhaustion overcame even this, the trains *awoke* us. (BNC misc.)

(101)  The falling bombs and strafing fire at Wheeler *woke* Lt. Kenneth Taylor. (CAE misc.)

(102)  “They knock on your door in the morning to *wake* you up and bring you tea”, she said. (CAE misc.)

(103)  There, it came again, just a quick flicker of sensation. “I've *woken* it up”, she thought, grinning. (CAE fic.)

To establish whether (103) really expresses meaning 05 or meaning 06, this example has to be discussed a bit more detailed. The *it* here refers to a foetus, and the questions that need to be answered is whether a foetus does sleep like a born human does, and if so, whether it really was sleeping the moment the mother thought she had woken it up. We cannot be sure about that, but the fact that the mother thinks she had woken her unborn baby from sleep is enough to claim that this example belongs to meaning 05 rather than 06.

Meaning 06, ‘to rouse someone from a state resembling sleep; to make active’, is a bit less frequent than meaning 05, but *awake* and *wake* are equally often used to express this sense. Fourteen instances were found for each verb. This meaning, similarly to its intransitive counterpart meaning 02, is frequently used in figurative senses. (104) – (107) below show how *awake* and *wake* are used in this sense. Examples (104) and (105) represent prototypical figurative uses of both verbs in the sense of meaning 06. The examples (106) and (107) then refer to figurative expressions involving the *sun* and the two verbs in question, exemplifying synonymy:
If Japan attacked the United States and **awoke** a sleeping giant, in a way so did we. (CAE spok.)

When the whole world is sleeping dead, a human being cries out and **wakes it up**; at that moment he creates history. (CAE acad.)

No ray of sunshine is ever lost, but the green which it **awakes** into existence needs time to sprout [...]. (CAE acad.)

Imagine the sun **waking** the flies to a confessional buzz in the camp latrines [...]. (BNC fic.)

Examples of a transitive use of *awake* and *wake* did not only occur in meanings 05 and 06. Two transitive instances occurred also within group 03:

These findings, coupled with his own desire to shatter the “reading is work” mentality that he had witnessed in the schools, impelled Trelease to take on what was fast becoming a crusade: **waking up** parents and educators to the importance of reading aloud - beyond the classroom. (CAE news.)

In spite of attempts to stop their transmission, the programmes went out on CCTV, watched by millions and caused a great deal of lively, heated debate. The complex, sometimes obscure series certainly was provocative, as it probably meant to be, **waking** the Chinese people **up to** harsh realities. (BNC acad.)

The two examples (108 and 109) show rare instances of meaning 03, basically expressing the sense ‘to awake someone to something; to make someone realise’. Those instances seem to be very rare. They were found in the academic and newspaper language genre, both of which tend to contain texts written by professional writers. Hence, it could be the case that the transitive form of meaning 03 is infrequent, because it is a very difficult expression grammatically (in terms of correctly inserting the object).

### 6.3 Summary of the qualitative analysis

This section summarises the findings of the qualitative analysis of *(a)rise* and *(a)wake*. Regarding the first verb pair *(a)rise*, we can see that the two verbs *arise* and *rise* basically complement each other. With a few exceptions (i.e. meaning 03), it was found that meanings that are infrequent for one verb are very frequent senses of the other verb instead. Most of the eighteen meanings shared by *arise* and *rise* are more usually expressed by means of [RISE] than [ARISE]. Additionally, *rise* is the only one of the two verbs that is used to express meanings other than those eighteen meanings shared by the two verbs. It can therefore be concluded that the semantic range of *rise* is wider than that of *arise*, and that the number of senses *rise* can express is higher than that of *arise*. Regarding the most frequent and most
typical meanings of the two verbs, we can see that both verbs are semantically specialised. [ARISE] is most frequent in those types of meanings that focus on the origin, a starting point or a cause of a movement (e.g. meaning 14, 17 and 18). [RISE] however is very frequently used when it comes to describing the movement itself (i.e. meaning 04, 07, and 11-13), an observation that was already made during the investigation of the collocates of the two verbs (see Section 5.7). As a conclusion for the qualitative analysis of those two verbs it can be said that, despite their similarity in their form and basic meaning, arise and rise show very distinct semantic values and are used in different circumstances. They can be synonymous in some senses, although other factors like style and context probably will determine whether a speaker chooses one or the other variant (especially when it comes to the distinction between poetic and formal language).

The members of the other verb pair (a)wake have a completely different relation to each other. The verbs awake and wake – in contrast to (a)rise – are frequently used in a synonymous way in most of the six meanings they share. The only difference was found in meaning 05, where wake seems to be far more common than awake. Similarly to [RISE], the verb wake was found expressing a meaning that cannot be expressed with awake, and the conclusion that [WAKE] has a broader semantic range than the verb awake might be supported. More such instances would be necessary however to enable us to draw an actual conclusion about this specific point.

The impact of transitive expressions with (a)wake is another difference to (a)rise that was identified. Meanings 05 and 06 are the transitive variants of meanings 01 and 02 respectively, the latter two of which are the most frequent uses of awake and wake. We can see that expressions of transitive meanings are more commonly realised by means of [WAKE]. It is therefore more common to say to wake someone from sleep or a state resembling sleep than to awake someone from such state, although the latter form is by no means uncommon. Concluding the qualitative analysis of the two verbs one can say that awake and wake are more synonymous than arise and rise, and that only slight differences in their meanings and semantic values could be identified. Basically, meanings that are common for one member of the verb pair are also common for the other one.

7. Discussion
This section will discuss the findings made in the quantitative and qualitative analysis and relate them to the theories that were presented as background information at the beginning of this paper.
7.1 Synonymy, semantic specialisation and semantic prosody

As already mentioned in Section 6.3, the qualitative analysis revealed that the constituents of the two verb pairs *rise* and *wake* have different relations to each other. In general, one can say that the members of both verb pairs have similar meanings. In terms of their use however, the constituents of *rise* and *wake* are different.

In very rare cases, *rise* and *rise* are frequently used in a clearly synonymous way (for instance in meaning 03). In more cases, their relation can be defined as propositional synonymy, referring to meanings in which *rise* and *rise* are paradigmatically exchangeable without changing the meaning of the expression as a whole (cf. meanings 05, 06, 08). In most cases, however, they are near-synonyms; stylistic and contextual factors favour the use of one and disfavour the use of the other verb. Meanings 01, 12, and 13 for instance are senses that are most frequently expressed with *rise* instead of *rise*. Meanings 17 and 18 on the other hand are more frequently expressed with *rise* instead. Since it is not impossible to express those meanings by means of the other verb pair constituent, and since the two verb pair constituents differ in their collocational and prosodic behaviour according to the different meanings they most frequently express, *rise* and *rise* can be regarded as near-synonyms in those cases.

Since there are cases where *rise* and *rise* are synonymous (however rare they are), the claim holds that in general there is a type of synonymous relation between the two verbs. To determine which of the three types of synonymy presented in Section 3.2 actually is the general meaning relation between *rise* and *rise* proves to be more difficult however. My conclusion is that they are near-synonyms. Many extra-linguistic (e.g. the personal preference of a speaker) and intra-linguistic factors (e.g. style and formality) define whether *rise* or *rise* should be used in a specific context. The collocational behaviour of both verbs coincides with the meanings the two verbs most frequently express: *rise* is most often used in economic contexts, describing the growth of *rates* or *prices* (i.e. meanings 12 and 13), whereas *rise* is bound to the growth or appearance of negative abstract concepts like that of *misunderstandings* and *difficulties* (cf. meanings 17 and 18).

Compared to *rise*, the constituents of the verb pair *wake* are far more synonymous. In many cases, they can be regarded absolute synonyms, as no differences in style or contextual factors could be determined. I am aware however of the fact that the concept of absolute synonymy is criticised by many linguists, and I would not want to go as far as claiming that *wake* and *wake* actually are absolute synonyms. The results of the investigations in this paper however indicate the existence of such a synonymous relation.
between the two verb pair constituents. They are most frequent in the same meanings (i.e. meanings 01 and 02), least frequent in the same meanings (i.e. meaning 04), and their frequency in the specific meaning groups are quite similar in general. Furthermore, they share certain collocational characteristics, and they are related to the same semantic fields, i.e. they have the same semantic specialisation (both are bound to the same concepts of ‘sleep’ and ‘time’ for instance). As a conclusion, awake and wake are more synonymous than arise and rise. Their synonymous relation is therefore above the level of propositional and near-synonymy. If the concept of absolute synonymy is ever accepted in linguistic research, the relation between awake and wake could probably be regarded an example of it.

Next to their synonymous relation to each other, the constituents of the two verb pairs differ in their semantic specialisations and their prosodic behaviour. Regarding the concept of semantic prosody, only arise seems to be closely connected to specific connotations, namely negative words (see Section 5.7). The other three verbs investigated favour neither positive nor negative words. Due to the observation that arise shows prosodic behaviour, and that awake does not, it seems as if this behaviour is not a specific characteristics of the prefix a- but of the verb itself.

7.2 (De)grammaticalisation, lexicalisation and unidirectionality

The two prefixed forms arise and awake are results of the language development process grammaticalisation. The ModE prefix a- in both forms probably has its origin in an OE preposition meaning ‘on’ (see Section 2.3 for a more detailed discussion of this rather problematic topic). In order for a form to become grammaticalised, it must undergo several other processes and fulfil specific characteristics beforehand. As Hopper and Traugott (2003: 99f) describe, “lexical items that become grammaticalized must first be semantically general and serve commonly needed discourse functions. They then become syntactically fixed (they become constructions), and may eventually amalgamate morphologically, say, as a stem and affix”. The outcome of the process described here is a frequently recognised diachronic tendency of language change, namely for analytic constructions to coalesce over time and become synthetic expressions (Hopper and Traugott 2003: 7f).

During this process, the meaning of the former OE preposition was bleached. This means that the number of its meanings decreased over time. A still unanswered question in linguistics is how far this semantic bleaching can go. Some linguists claim that a total loss of the meaning of a grammaticalised form can occur, such as Lehmann (1995: 129), who claims that “[g]rammaticalization rips off the lexical features until only the grammatical features are left”. Hopper and Traugott (2003: 95f) however say that usually “some traces of its original
lexical meanings tend to adhere to it, and details of its lexical history may be reflected in constraints on its grammatical distribution”. Furthermore, it is not unusual for synthetic and analytic expressions to coexist and thereby express synonymous meanings (cf. Hopper & Traugott 2003: 7f). Regarding the verb arise, this observation seems suitable: the verb is quite frequent, although it coexists with rise, which, in combination with prepositions (mostly up), can be used to express many meanings of arise in a more or less synonymous way. This seems to be due to the specific characteristics of the prefix a-, characteristics that add semantic specialisations to this verb which motivate the speakers to prefer it over periphrastic variants with rise. Such characteristics that were found for the verb arise are, for instance, the focus on the origin or starting point of a movement or process and the specialisations on certain semantic fields. For the verb awake however, this is not really the case. It is very infrequent and it is used to express meanings that are very similar to those expressed by the verb wake in combination with prepositions. It seems as if the prefix a- in awake is not as semantically specialised as it is the case for arise. It could be the case that the process of morphologisation, as part of the general process of grammaticalisation, had different impacts on the characteristics on the prefix a- in the two verbs arise and awake.

According to many theorists, grammaticalisation is a unidirectional process. Hopper and Traugott (2003: 131f) explain that “[t]he strongest [claim] is that all grammaticalization involves shifts in specific linguistic contexts from lexical item to grammatical item, or from less to more grammatical item, and that grammatical clines are basically irreversible”. In Section 3.3.2, mechanisms were presented that are part of the grammaticalisation process. One of those mechanisms is called renewal, or renovation. In Section 5 and 6 it was shown that many meanings of arise and awake can also be realised with help of their shorter variants rise and wake and the preposition up. The use of those periphrastic expressions can be interpreted as a renewal of the prefix a- in the verbs arise and awake. Given the similarity in meaning between the prefix a- and the preposition up (in Section 2.3 it was shown that ‘up’ actually is one of the meanings of a-), it is not really unexpected that speakers realised this similarity and – especially in times of coexistence of the forms arise and rise and awake and wake – started to replace the synthetic expression with an analytic one. By replacing the prefix a- with a preposition with the meaning ‘up’, rise up and wake up are two forms that can be recognised as results of a process that makes grammatical items into a lexical one. According to Hopper and Traugott (2003: 8), this process is as usual in language change as the grammaticalisation process that turns lexical items into grammatical ones: “The second diachronic tendency […] is an example of what is known as ‘renewal’ – the tendency for
periphrastic forms to replace morphological ones over time”. An interesting question arises (taken from Lehmann 1995: 19), which refers to a highly debated topic within grammaticalisation theory: “If such changes from the synthetic to the analytic do occur, aren’t they instances of degrammaticalization?” In view of the theoretical background, the answer is yes: grammaticalisation is the change of a lexical item into a grammatical one. The inversion of this definition would then refer to the process of degrammaticalisation instead, the acknowledgement of which would question the unidirectionality of grammaticalisation in general.

With regard to the model of grammaticalisation theory however, the answer is not as easy. In his discussion about this topic, Lehmann (1995: 3) refers back to the early days of grammaticalisation theory, by quoting the 19\textsuperscript{th} century linguist Georg von der Gabelentz\textsuperscript{5}, who already in the 19\textsuperscript{th} century noticed that language change and development is neither linear nor cyclic, but that it, driven by the two needs of *ease of articulation* and *distinctness of expression*, can be explained by means of a spiral. If grammaticalisation was a purely linear process, the languages of the world would run out of lexical items, as they all would have been turned into grammatical items at one point. On the other hand, it cannot be circular, because that would mean forms are grammaticalised just to be degrammaticalised again, turning them into the same lexical items that were the starting point of the grammaticalisation process in the first place. Here, the process of *renewal*, or *renovation*, plays an important role. Lehmann (1995: 20) explains that “[i]f the evolution along grammaticalization scales takes the form of a spiral, this implies that forms which are given up near the end of the scale may be substituted by new forms entering at its beginning”. The necessity of this renewal process has been described by Lindström (2004: 82): “[B]ecause expressions will be economised until they can hardly be recognised and stop to fill their function, they must at some point be reinforced for clarity, and then the new expression will start its path towards ‘a better economy’”.

\textsuperscript{5} “Nun bewegt sich die Geschichte der Sprachen in der Diagonale zweier Kräfte: des Bequemlichkeitstriebes, der zur Abnutzung der Lauten führt, und des Deutlichkeitstriebes, der jene Abnutzung nicht zur Zerstörung der Sprache ausarten läßt. Die Affixe verschleifen sich, verschwinden am Ende spurlos; ihre Funktionen aber oder ähnliche drängen wieder nach Ausdruck. Diesen Ausdruck erhalten sie, nach der Methode der isolierenden Sprachen, durch Wortstellung oder verdeutlichende Wörter. Letztere unterliegen wiederum mit der Zeit dem Agglutinationsprozesse, dem Verschleifen und Schwunde, und derweile bereitet sich für das Verderbende neuer Ersatz vor: periphrastische Ausdrücke werden bevorzugt; mögen sie syntaktische Gefüge oder wahre Komposita sein (englisch: I shall see, - lateinisch videbo = vide-fue); immer gilt das Gleiche: die Entwicklungslinie krümmt sich zurück nach der Seite der Isolation, nicht in die alte Bahn, sondern in eine annähernd parallele. Darum vergleiche ich sie der Spirale”. (von der Gabelentz 1901: 256, quoted in Lehmann 1995: 3)
Now, especially in view of the verb *rise*, a tendency can be observed that shows a further development of the renewal process described above. The forms *uprise* (used as noun and verb) and *uprising* (used as noun or participial adjective) could be interpreted as a derivation of the preposition *up* and the stems *rise* and *rising* respectively. The results are these prefixed forms, and the process involved here is a quasi-repetition of the development of the OE preposition ā and the verb rísan to ārísan, which was already defined as grammaticalisation. The derivation of the preposition *up* and the forms *rise* and *rising* might have been a result of a process in which the older and less productive form a- was substituted by the more modern and possibly more common form up-, thereby exemplifying the renewal process mentioned above. Therefore, one could regard it as part of the spiral grammaticalisation process as it has been described here. One question arises however: when did this change occur?

Unfortunately, the OED gives only little information about the etymology of the forms *uprise* and *uprising*. The only information we can get is that the earliest records of these lexemes are from the 13th century. It seems as if this development took place already some 800 years ago.

One major difference between *uprise* and *uprising* suggests another conclusion however. On the one hand, the form *uprise* is almost obsolete in ModE. In each of the two corpora BNC and CAE, only one instance could be found, used as verb in the former and as noun in the latter corpus. On the other hand, the form *uprising* is still very prominent in ModE. It has a frequency of 3.8 instances per one million words in the BNC, and in the CAE it occurs even 5 times in one million words. It could be the case that, over the centuries, the speech community forgot about the forms *uprise* and *uprising*, because there were synonymous and frequent expressions like *arise* and *arising*. In present times, in which the prefix a- is not very productive anymore, speakers probably have problems to identify the semantic values of this prefix. Regarding *arising*, the renewal process might have taken place again, resulting in the form *uprising* that is frequently used today.

However, given the existence of a (in most senses) synonymous and frequent form *arising*, the frequency and importance especially of the form *uprising* – whether it is used as an adjective or a noun – is unexpected. Especially in combination with the verb *rise*, the prefix a- seems to have specific semantic characteristics. Hence there is no need for speakers to renew the forms involved in the verb *arise*. My assumption is that the meaning of the prefix up- is clearer than that of the prefix a-. In ModE, the prefix a- is not very productive anymore. It might therefore be easier for speakers to use the prefix up- instead, promoting the use of *uprising* and driving out the use of *arising* respectively.
Although the OED claims the existence of an old and rarely used form *upwake* (used as a verb), no instances of this form have been found in the two corpora used in this study. It can therefore be concluded that, to the best of my knowledge, no tendency that is comparable to the renewal process of *arise* has occurred for *awake* yet, although, in view of the characteristics of this verb (e.g. its high degree of synonymy with *wake*), such a change would be motivated to a greater extent. There might be similar cases among other word pairs that are distinguished by the presence or absence of the prefix *a*- however. Due to the lack of space and time, the actual relation towards and difference in frequency between those two lexemes cannot be investigated here. It might be a suitable topic for other papers to concentrate on the terms *arising* and *uprising* or similar expressions.

Although the topic of degrammaticalisation is hotly debated in linguistics, and although many grammaticalisation theorists doubt the existence of such a process, I want to go that far and claim that the above described cline of *a- + verb > verb + up (> up- + verb)* involves a case of degrammaticalisation in terms of splitting the synthetic form *a- + rise/wake* into the two lexical items *rise/wake + up*. The renewal process described here is a degrammaticalisation process, because the ‘new’ form *up* does not influence the lexeme itself, as the verb and the preposition form a periphrastic expression. On the morphological level, through the process of degrammaticalisation, a synthetic expression is turned into an analytic one. Since the process of turning a grammatical form (e.g. a prefix) into a lexical form (e.g. a preposition) moves into the opposite direction of grammaticalisation (a process often regarded as being unidirectional), solely functionally, it is a form of degrammaticalisation. Admittedly, this phenomenon can also be classified as lexicalisation. The boundary between the two processes is very small however, as van der Auwera summarises in a somewhat oversimplified way, as he admits:

> The undoing of grammar invites the term ‘degrammaticalization’ and the making of lexicon invites the term ‘lexicalization.’ But there is only one process: with ‘degrammaticalization’ one looks at it from one end, and with ‘lexicalization’ from the other. With the same reasoning, ‘delexicalization’ could be coextensive with ‘grammaticalization’.

(van der Auwera 2002: 20)

However, at the moment when a preposition is turned into a prefix this process has to be regarded as going into the direction of grammaticalisation again. Since in the word *uprising*, the preposition *up*, the renewal of the old prefix *a-*, was turned into a prefix, this development has to be interpreted as renewal or renovation instead. The newer form *up-* takes the place of the older variant *a-*, and as a prefix, it directly influences the characteristics of the lexeme.
Therefore this development has to be regarded as part of the spirally formed grammaticalisation process. Hence, in view of the development of arising to uprising, the degrammaticalisation process can be seen as part of the general grammaticalisation process, being responsible for the renewal process that leads to the further development of the preposition up into a prefix. “In this sense”, as Lehmann (2002: 1) concludes, “grammaticalization presupposes lexicalization”.

As a general conclusion, the argument above is not suitable, however. It might be the case for arise that degrammaticalisation and the renewal process are part of a further grammaticalisation process, in view of (a)wake however this is not the case. The only development that could be found was the degrammaticalisation, or lexicalisation, of a- + wake into wake + up. Lehmann comments on such developments, stating that a case where “a new preposition or conjunction evolves […] is not a case of grammaticalization, because the particle thus developed is not a grammatical element”. As we are still waiting for a renewal process of the verb awake that is analogous to that of arise, resulting in a frequently used form *upwaking or similar (cf. analogy in Section 3.3.2), the developments for arise and awake could be interpreted as two different processes. Given the similarity between the forms and the characteristics of the two verb pairs (a)rise and (a)wake however, this is highly unexpected, as the question arises why two forms involving the same constituents (i.e. the prefix a- and verb stems) would develop into different directions. Below, a short discussion about the role of the prefix a- in the developments presented in this section will end the discussion about the findings of this study in view of the theoretical background.

7.3 The role of the prefix a-
Interestingly, similarly to the ambiguous observations made in the discussion about the origin and meaning of the prefix a- in Section 2.3, the findings of this study do not really allow far-reaching conclusions about the general role of this prefix in the observations described in this study. On the one hand, in the investigations above, observations were made that indicate a certain similarity between the two verbs with the prefix a-. They are both (more or less) synonymous to their unprefixed variants (often in combination with the preposition up), a factor that, especially in terms of grammaticalisation and degrammaticalisation, can be referred to the characteristics of the prefix a- (see above). Furthermore, arise and awake show certain similarities; in sentences like “Awake and Arise!””, taken from poetic, religious or simply literal contexts, the two verbs complement each other. They are semantically related as they can refer to the same semantic concept of ‘sleep or states resembling sleep’; arise as well as awake can mean ‘to get out of sleep or a state resembling sleep’ and ‘to realise something
and rise from inaction’. This however is also true for rise and wake and cannot really be traced back to the character of the prefix a-.

On the other hand, the verbs arise and awake are different in their relation to their verb pair constituents rise and wake respectively. In both cases, the prefix a- adds the meaning ‘up’ to the stems, but in case of arise, this process has a deeper impact on the semantic value of the verb than it is the case for awake. The former verb is frequently used to express a focus on a starting point or an origin of a movement. Whether this specialisation is due to a characteristic of the prefix a- or a simple semantic specialisation of the verb arise is not clear. The fact that such an observation could not be detected for awake would support the latter point however.

If this semantic specialisation was a consequence of a certain characteristic of the prefix a- in arise, it would become clear that we actually are dealing with two different prefixes in the verbs arise and awake, a possibility that has already been discussed in Section 2.3. It could also be the case that speakers do not recognise the similarity of the prefixes in arise and awake, awarding specific semantic values to the former, but not to the latter verb. They seem to no longer recognise any specific semantic value that would motivate them to use awake more frequently.

Can we conclude that the prefix a- in awake lost most of its meanings that distinguished it from its verb pair constituent wake, while in arise it did not? This would explain the infrequent use of awake observed in both corpora. The verbs awake and wake are very synonymous, and, according to Aitchison (2001:178), “language tends to eliminate pointless variety and prefers constructions which are clear and straightforward”. Eliminating the rather unclear form involving the more or less unproductive prefix a- seems to be a natural development, and quite an old one, as Kastovsky points out:

[K]he system of OE prefixes, in particular those occurring with verbs, was already at the end of the tenth century in a state of advanced decay, because many prefix-verb combinations were no longer transparent. With many verbal prefixes, e.g. a-, ge-, ob-, it is impossible to establish consistent meanings, and frequently there does not seem to be any meaning difference at all between the simplex and the prefixed form. This is confirmed by the observation that in subsequent copies of one and the same text prefixes are often omitted, added or exchanged for other prefixes without any apparent semantic effect […]. This points to a considerable weakening of the meaning of these prefixes, especially of a-, be-, ge-, […].

(Kastovsky 2005: 377)
Once again, this could mean that we are dealing with two different prefixes, since in arise, the prefix carries more semantic values, which makes the differentiation between arise and rise important and both verbs quite frequently used. Scholars have problems in identifying the exact meanings of the prefix a-. Kastovsky (2005: 378) writes that “In view of the vagueness of the meaning of this prefix [a-], [...], it is difficult to give precise semantic patterns [...] In some instances it seems to denote ‘out’; in others, it seems to add an intensifying [...] element [...]. But in the overwhelming majority of instances, its meaning is no longer transparent”. Those different semantic values of the prefix a- could be referred back to the different origins this prefix is thought to have. It does not seem too unlikely that arise and awake are derivatives with prefixes from two different origins. If we are actually dealing with two different prefixes however, the question arises why many meanings of awake can be expressed by means of the unprefixed verb wake in combination with the preposition up – a feature that also is characteristic for the verb pair (a)rise. It seems to be necessary to investigate even more word pairs distinguished by the presence or absence of the prefix a- to get a clearer picture of this.

8. Conclusion

In this section, an attempt is made to answer the seven research questions that were posed at the beginning of this paper. Furthermore, suggestions for further research will be made in order to motivate additional investigations on this topic.

Answers to question a), “What meanings are most frequently expressed by arise and rise and awake and wake?”, were found in the qualitative analysis. It was observed that arise and rise are most frequent in different meanings. The most frequent, and probably most typical meanings of arise are meaning 17 (i.e. ‘to result from’) and meaning 18 (i.e. ‘to spring up’). The verb rise is most frequent in the meanings 13 (i.e. ‘to come up to a point in scale’) and 12 (i.e. ‘to rise in price or amount’). Regarding (a)wake, the verb pair constituents behave differently from (a)rise. The verbs awake and wake are most frequent in the same meanings, such as meaning 01 (i.e. ‘to come out of sleep’).

Regarding question b), “Can any sort of semantic specialisation be detected?”, it can be summarised that specific semantic specialisations could only be detected for the verb pair constituents of (a)rise. The two verbs awake and wake are closely bound to the semantic field of ‘sleep’ and to temporal expressions like morning and dawn, all of which can easily be related to the semantic field of ‘sleep’ as well. The verbs arise and rise show different semantic characteristics. As the quantitative analysis revealed, arise very frequently occurs
with abstract nouns having negative connotations. It could be the case that arise itself has developed a semantic prosody, being recognised as a negative verb in general. Furthermore, it was found that there is a difference in focus between expressions with arise and rise. When using the former verb, the focus is on an origin or a starting point of a movement, whereas expressions with rise tend to focus solely on the manner of a movement. Furthermore, in contrast to the characteristics of arise, the verb rise is very frequently used within the semantic fields of ‘money’ and ‘economy’.

Question c), “Are arise and rise or awake and wake synonyms? What type of synonymy is there?”, can be answered only vaguely for (a)rise; for (a)wake however the relation between the two verb pair constituents seems to be clearer. The verbs arise and rise, despite being most frequent in different meanings, can, under certain circumstances, be used in a synonymous way. It was decided to define the relation between those two verbs as near-synonymy. The verbs awake and wake are far more synonymous. As has been discussed above, characteristics of absolute synonymy were found in many meanings shared by those two verbs. However, the conclusion that the two verbs are absolute synonyms is debatable, as is the concept of absolute synonymy in general.

One observation that was made in the qualitative analysis can be related to question d), “Can any other semantic relation between arise and rise or awake and wake be detected?”. Regarding the verb pair (a)rise, a relation between the two constituents was found that can be interpreted as complementary: the two verbs, despite their near-synonymous relation, have specialised in different meanings. The infrequent use of one verb in a specific meaning category is complemented by the frequent use of the other verb pair constituent expressing this meaning category. Regarding awake and wake, no such relation was found.

Question e), “Is the assumption (made by the OED), that the verb arise is being supplanted by rise, detectable in language use? Does the same apply for awake and wake?”, relates to the different characteristics that were found for the verb pair constituents of (a)rise and (a)wake. Regarding arise, it can be concluded that it is a rather stable verb, being used most frequently in meanings that are not really typical for rise. Due to the fact that rise is infrequent in many of the meanings that are not really typical for rise. Due to the fact that rise is infrequent in many of the meanings that are temporarily expressed with arise, and due to the rather high frequency of the verb arise in the two corpora investigated for this study, the assumption that arise is being supplanted by rise does not hold. It has to be said however that the verb rise is far more frequent than arise. This difference in frequency – as well as the near-synonymous relation between the two verbs – might indicate that the latter verb is going to be supplanted in due time. Regarding the verb pair (a)wake however, this process seems to
be far more advanced. The findings of this study indicate that the verb *awake* is being supplanted by *wake*: *awake* in general is very infrequent in the two corpora investigated, and taking into account the high degree of synonymy between the two verbs, it can be concluded that *wake* is replacing *awake* in language use.

Question f), “What are the reasons for those tendencies?”, relates to the reasons for the developments discussed in the paragraph above. The main reason for the observation that one verb pair constituent is being supplanted by the other one seems to be a high degree of synonymy between the two words. Aitchison (2001:178) claims that “language tends to eliminate pointless variety and prefers constructions which are clear and straightforward”. Especially with regard to the verb pair *(a)wake*, the two constituents of which are highly synonymous, one verb is regarded as pointless and unclear by the speakers. The logical outcome is that speakers do not use this verb anymore, preferring the other one instead.

That it is the verb *awake* which is on the verge of disappearing can be explained by the characteristics of the prefix *a-*-, which is rather unproductive today. Furthermore, as discussed in Section 7, the origin and even the general meaning of this prefix are unclear, even to linguists. It seems likely that speakers have problems recognising the semantic values of the prefix *a-* and prefer to use the clearer unprefixed variant *wake* instead. Unfortunately, such explanations can only be related to *(a)wake* however, since the verb pair *(a)rise* does not show such tendential developments (or at least not to the same extent). We might be forced to settle for lexeme specific explanations rather than for generalisations.

On the other hand, in relation to processes of semantic change that are part of grammaticalisation theory, *(a)rise* and *(a)wake* show similar tendencies. Many meanings of the analytic forms *arise* and *awake* tend to be expressed with synthetic expressions like *rise up* and *wake up* respectively, describing the process of *renewal*, a general tendency of language change according to Hopper and Traugott (2003: 8).

According to such ambiguous findings and vague explanations as described above, the final question g), “What role does the prefix *a-* generally play in the tendencies observed? Can we generalise about the role of the prefix *a-*?“, cannot really be answered. As already discussed in Section 2.3, it is very difficult to get a general picture of the characteristics of the prefix *a-* involved in the creation of the verbs *arise* and *awake*. Too many behavioural differences between the two verbs could be found, some of which might be due to the influence of the prefix *a-*-, while some of them might not. On the other hand, there is one major similarity between *arise* and *awake*: a renewal process is leading to the tendency that many meanings of the two verbs are expressed by means of periphrastic phrases like *rise up*
and wake up. We cannot be sure about the role the prefix a- plays in the tendencies and developments discussed above. Therefore it is difficult to generalise over the prefix a-. Discussions throughout this study showed that we cannot even be sure whether it is only one prefix or if there are two different prefixes involved in the derivations arise and awake. Many findings indicate that the latter assumption is correct. It seems to be necessary to investigate more word pairs that are distinguished by the presence of absence of the prefix a- to get a more general picture of the characteristics of this prefix and of its role in language change processes.
References

Primary sources

Secondary sources


