

Phrasal Settings in Which the Definite and Indefinite Articles Appear to Be Interchangeable in English: An Exploratory Study

Stephen James Coffey  

University of Pisa, Pisa, Italy
stephen.james.coffey@unipi.it

Abstract. This paper investigates the fact that some English phrases may be used with either the definite or the indefinite article, without any obvious change in meaning or effect. Whereas many types of phrasal variation have been documented in the literature, there is little description of this particular phenomenon. Examples are given of a number of phrasal types, ranging from opaque idiomatic expressions to fairly transparent word strings. Contextualized examples are provided, mainly from the British National Corpus but also from corpus-based reference works.

Keywords: Definite article · Indefinite article · Esphoric NPs · Corpora · Phraseology

1 Introduction

The words *the* and *a* (or prevocalic *an*) are usually mutually exclusive in English text; this is only natural since they have different uses, for the most part; indeed, their functions are often in direct contrast to one another. One area in which they do, partially, overlap, is that of ‘generic reference’, and here the two words are sometimes interchangeable. Consider, for example, the following extract, from which an article has been deliberately removed and where I would judge both articles to be possible at that point in the text:

A buzzard’s feathers have almost a glossy look to them, and they’re more than one shade of brown. If you observe the bird from the front, you’ll see that it ranges in hue from the creamy white of the breast and stomach to quite a dark brown with even darker flecks on the neck. When ___ buzzard is in flight, cream is the main colour you see, and if you have good eyesight and know what you are looking for, it’s quite distinctive, even from a distance.

[This extract was retrieved from the British National Corpus, where the reader may also check which article was used in the original text.]

The two possibilities in this text correspond to two of five patterns for ‘generic reference’ listed in Berry [1]; these are, quite simply, ‘the indefinite article + singular count noun’ and ‘the definite article + singular count noun’ (p. 35). The example sentences provided are, respectively, ‘A dog likes to eat far more than a human being’ and ‘The gorilla is a shy retiring creature.’ Berry, however, warns against assuming that the various patterns for generic reference are always interchangeable.

As well as being occasional alternatives before single-word units (as with *buzzard* in the above extract), *the* and *a/an* can also be found as alternatives within phrasal units of one sort or another, and it is this aspect of their usage which is the focus of the present paper. In Sect. 2, I exemplify and discuss *specific* phrases in which either article may be found, while in Sect. 3, examples are given of relevant phrases belonging to the lexicogrammatical frame ‘*a/the N₁ of N₂*’. The emphasis in the study is more on British English than other geographical varieties, since the corpus used to investigate phraseology was the British National Corpus; also, the dictionaries consulted, even though including data from the US and other countries, were published by UK-based publishers.

2 Specific Phrases

As is well known, many fixed phrases are only *relatively* fixed. One of the main variation types is that of alternative component words, as in *a piece of the action/a slice of the action*, usually described as two versions of what is essentially the same phraseological item. Many examples of such variation are given in the literature, and in relation to different word types. The following phrasal pairs, for example, include alternative particles and conjunctions: *a bolt from the blue/a bolt out of the blue*; *hit and miss/hit or miss*; these examples are taken from [2] (p. 129). Article variation, by contrast, has received little attention; an example from my own data is *do sb a world of good/do sb the world of good*.

In order to help presentation, and also to reflect the methodology of the study, examples of article variation are divided into two broad groups. Firstly, I discuss phrases which are to some degree semantically opaque when considered from the perspective of their component words, and secondly I give examples of phrases which are reasonably transparent from this point of view. In both cases the article within a phrase is being considered as an integral part of that phrase. To give a clear example of what I mean by this (though not involving article variation), the word *a* is a necessary component of the verbal phrase *come to a head*, and the word *the* is always found in the phrase which has as its lexical nucleus the words *the naked eye* – the presence of the articles in these two particular phrases is commented on by John Sinclair in, respectively, [3] (p. 161) and [4] (pp. 83–89).

2.1 Semantically Opaque, or Partially Opaque, Phrases

One of the sources of phrases in this study was corpus-based dictionaries of idioms. This was considered to be a useful starting point since lexical variation is well attested in the types of phrase typically included in such dictionaries. To this end, searches were carried out in two dictionaries: the *Collins COBUILD Dictionary of Idioms* [5], hereafter CCDI, and the *Cambridge International Dictionary of Idioms* [6], CIDI. According to descriptive information in the dictionaries themselves, the Cobuild dictionary includes ‘approximately 4400 current British and American idioms’, while CIDI includes ‘around 7,000 idioms’ covering ‘current British, American and Australian idioms’. Both dictionaries were compiled mainly for learners and teachers of English as a foreign language.

Dictionary consultation involved looking at not only the citation forms of phrases, but also examples of usage. This was because sometimes variation was found only in examples, or through a combination of citation form and examples; in the case of the Cobuild dictionary the definitions were also of relevance. Examples of usage come from, or were based on, the corpora used to help compile the dictionaries, respectively, the Bank of English and the Cambridge International Corpus.

The phrases found in either or both of these dictionaries form a heterogeneous set from the point of their semantic composition. They include, among others, phrases with more literal counterparts (e.g. *come to a boil*), similes (*as dead as a dodo*), phrases in which one or more words are being used in their usual sense (e.g. *not a ghost of a chance*), and phrases in which a specific sense of a word is associated with a specific structure (e.g. *a model of [politeness, etc.]*).

Not too many relevant examples of article variation were found in the two idioms dictionaries. This may simply reflect the fact that there is not very much variation of this sort in the English language. Another consideration, however, is the fact that in some phrases one of the two alternative forms may be found much more frequently than the other, and for this reason the less frequent form may not have been included in the dictionary, both for reasons of space and so as to keep presentation relatively simple for the dictionary reader. A further factor is the relatively low corpus frequency of many phrases of this sort, and consequent lack of available data (see [7], pp. 311–316, [8], pp. 82–90).

In addition to relevant data found in the two dictionaries, a list was also made of phrases which were judged to be possible candidates for variation of the sort being investigated. These were then looked for in the British National Corpus (BNC), as well as in general corpus-based dictionaries. The BNC was consulted using the University of Lancaster's BNCweb interface (<http://bncweb.lancs.ac.uk/>).

Before giving examples of the phrases found, it should be pointed out that in all cases, the articles are being considered as genuine alternatives and the choice of one article or the other is not related to the precise nature of the context. This is very different from cases in which the usual form of a phrase has been adjusted in some way so as to fit into a specific context. Noun modification is quite common in this respect. Here is an example, involving the phrase *the thin end of the wedge*, followed by a non-modified example of the same phrase:

'I understand the intention – of wanting to try to make a conciliatory gesture – but this could be the thin end of a very dangerous wedge.' [BNC]

'Opponents say legalisation would be the thin end of the wedge.' [BNC]

A further example of modification, though slightly different in nature, is the following: 'But I don't speak French or Arabic so I'm going to stick out like the proverbial sore thumb.' [BNC]. Here, the phrase *stick out like a sore thumb* has been adapted so that the expression is not only being used, but is also being referred to metalinguistically through the words 'the proverbial'. One could, perhaps, also have found '... like a proverbial sore thumb', but, on the basis of evidence in the BNC, reference to idioms and sayings through use of the word *proverbial* almost always involves the word *the* rather than *a*. The same findings are reported in [2] (pp. 306–7) relating to other corpora.

Some examples of ‘genuine’ article variation which were found in one or both idiom dictionaries are: *as dead as a/the dodo*, *come/bring to a/the boil*, *jump in a/the lake*, *to play a/the waiting game*, *given half a/the chance*, *not a/the ghost of a chance*, and *a/the model of*. I now discuss each of these pairs in turn.

The first expression has as its citation form in CCDI the form *dead as a dodo*. The definition, however, is more precise regarding form: ‘If you say that something is as **dead as a dodo** or as **dead as the dodo**, you mean that it is no longer active or popular’; at the same entry there is also an example of usage for each of the two forms:

‘The foreign exchange market was as dead as a dodo.’

‘This lugubrious Mozart style is as dead as the dodo everywhere in the world except Vienna and Salzburg.’

The next phrasal pair has the following explanation in CCDI: ‘If a situation or feeling **comes to the boil** or **comes to a boil**, it reaches a climax or becomes very active and intense’; ‘Someone or something can also **bring** a situation or feeling **to the boil** or **bring it to a boil**’ (pp. 39–40). The two examples given are with the verb *come*, and a different article is used in each case:

‘Their anger with France came to the boil last week when they officially protested at what they saw as a French media campaign against them.’

‘The issue has come to a boil in Newark, where federal prosecutors have warned lawyers that if the chairman is indicted, the government may move to seize the money that he is using to pay legal fees.’

The next two phrases, *Go jump in a/the lake* and *to play a/the waiting game*, are presented with article variation in CIDI, though there is only one example of usage in each case. Two examples of the second phrase, taken from the BNC and with different articles, are the following:

‘There is no doubt about Mary of Guise’s political ability. She had played a waiting game with great skill in the 1540s’

‘Think before saying yes — or no. Consider the value of a family contract. Prepare to play the waiting game.’

By contrast, in the following example the presence of a prenominal modifier appears to exclude the possibility of *the*, or at least to make it sound more awkward:

‘PETER Coyne is 80 min away from a Sydney Grand Final appearance — leaving Castleford to play a frustrating waiting game.’

The phrase *given half a/the chance* is presented in CIDI in precisely that form, that is with either article, and the same is true for some learner’s dictionaries. In the BNC, *given half a chance* is by far the more frequent form. Examples with the two respective articles are:

‘It’s the trees they go for, given half a chance.’

‘I am always happy to work myself up into a great cultural stew, given half the chance.’

Actually, phraseological description is more complex than this, especially since the slightly shorter phrase *half a/the chance* is also found in *if*-clauses, notably with the verbs *give* or *get*:

‘Not in an evil way at all, but if you gave him half a chance he’d hammer you into the ground and stamp on you.’

‘All women take advantage of the men in their lives if they get half a chance.’

The only example found in the BNC of ‘half *the* chance’ in an *if* clause is with the verb *have*: ‘I never want nothing on the side’ – ‘You could if you have half the chance’ (transcription of a conversation).

A second phrase with the word *chance* is *NOT a ghost of a chance*, where *NOT* indicates any negative word. The citation form in CIDI is with *a* (and not *the*), though the only example given is with initial *the*: ‘Against competition like that, they didn’t have the ghost of a chance of winning’. There were no examples with *the* found in the BNC, but there again, there were only five tokens with *a*; an example is: ‘You would have thought *The Woman In Black* wouldn’t have stood a ghost of a chance of survival with just two actors, a simple set and some offstage sound effects going for it.’

The next phrase exemplified is *a/the model of*. It is presented with both articles in CIDI, though there is just one example, with the word *the*: ‘Claudia, always the model of good taste, looked elegant in a black silk gown’. Judging from dictionary presentation generally, as well as evidence in the BNC, the phrase appears to be used above all with the word *a*. Examples of both forms from the BNC are:

‘Apart from one or two lapses, you’ve been a model of fidelity.’

‘Rangers remain the model of consistency, aiming tonight for their 38th successive match without defeat.’

Generally, this phrase implies approval on the part of the writer or speaker, though not always: ‘The Whitehall switchboard was a model of inefficiency, as usual’.

I turn now to phrases which were not found with alternative forms in either dictionary but which were among those found in the BNC. Examples are the following, and in most cases one of the articles (in brackets) was found to be much less frequent than the other: *put a (the) damper/dampener on STH*, *give SB the (a) cold shoulder*, *talk the (a) hind leg(s) off a donkey*, *the (a) lion’s share*, *at the (a) drop of a hat*, *a/the ghost of a smile*, *do SB a/the world of good*, and *a (the) means to an end*. The ‘hind leg’ idiom is interesting, compositionally, in that the only logical possibilities should be *talk the hind legs off a donkey* and *talk a hind leg off a donkey*; however, the form *talk the hind leg off a donkey* is also found (and, indeed, is the citation form in [9]), despite the fact that this wording suggests that donkeys only have one hind leg. The following are short contextualizations for some of the relative corpus examples:

‘... it might put a damper on things if you remained tight-lipped and poker-faced.’

‘So that put the damper on gigs for a bit.’

‘Both Price Waterhouse and now Touche Ross have dallied with the Deloitte domestic partnership and been given the cold shoulder.’

‘Where Italy is being given a cold shoulder today, could other deficit countries with particular political problems [...] be jilted tomorrow?’

‘Someone who’ll argue the hind leg off a donkey just for the sake of it.’

‘Aye, and such a one’ – she was shaking her head – ‘who’d talk a hind leg off a donkey ...’

‘They fear that having separated post and telecommunications, France Telecom will receive the lion’s share of the budget and may ultimately be privatised altogether.’

‘Rupert Murdoch’s TV Guide could be in the strongest position, to scoop a lion’s share of the market, since it is already established.’

‘He’ll give hundreds away at the drop of a hat.’

‘Mr Rodger Bell QC, for Mr Bewick, suggested that his client had a “bee in his bonnet” about surgeons being able to work anywhere at a drop of a hat.’

The phrase *a/the ghost of a smile* was not included in either idioms dictionary, and was investigated because of its similarity to *a/the ghost of a chance* (though without the negative patterning of the latter). It was found with both articles in the BNC, five tokens with *a* and seven with *the* (ignoring one with a prenominal modifier). Examples are:

‘He bowed his head a fraction, a ghost of a smile on his mouth.’

‘The ghost of a smile glimmered in his eyes.’

The next phrase, *do sb a/the world of good*, is presented in a number of dictionaries in just this way (i.e. with either article), and evidence from the BNC suggests that both forms are used indifferently. There are 31 tokens in all in the BNC, 14 with *a* and 17 with *the*. Examples with each article are:

‘... this type of exercise will do you the world of good.’

‘A good run in pastures new would do you a world of good.’

The last of the phrases listed above is *a/the means to an end*. In this case the preferred form, both in dictionaries and in the corpus consulted, is clearly the version with *a*. In the BNC there are 65 examples of *a means to an end* and six of *the means to an end*. Examples of the two forms are:

‘All this could not have been achieved without Macintosh computers, but again, they were only the means to an end.’

‘There is considerable debate about the value of education for its own sake, rather than as a means to an end.’

Lastly in this section I will mention two pairs of related phrases. The first, included in CIDI, is *one way or the other/one way or another*, and the second is *one after another/one after the other*. Of course, these phrases do not contain the word *an*, at least not as a separately written word, so strictly speaking they should not form part of this study. However, it should not be forgotten that there is no a priori reason why the phrase *an other* should not have become the norm in modern English; compare Sinclair’s comments regarding the compound nature of the ‘phrase’ *of course* and the ‘words’ *maybe, anyway* and *another* ([10], p. 321).

The expression with the word *way* is accompanied by one example sentence in CIDI: ‘One way or the other, I’m going to finish this job next week.’ An example with *another* from the BNC is: ‘One way or another, I’ll get into the garden. So it doesn’t matter what happens!’ It should be noted, however, that the two phrases can be used in different ways, and are not always interchangeable. For example, in the following extract it would not have been possible to use the form with *another*. ‘Claims that it shortened the war by several years cannot be proved one way or the other’. For more discussion of the phraseological nature of the word *way*, see [3].

The other pair of phrases involving ‘other’ are both used quite frequently in English, and appear to be interchangeable in most circumstances. Examples are:

‘The usual pattern is that you take several fish one after the other, and then nothing for perhaps an hour...’

‘It is depressing to see the predicted problems materialising one after another...’

However, if we allow for an intervening word after the word *one*, which usually means a noun, the comparison between the phrases alters somewhat. The BNC frequency figures for the basic phrases are: *one after the other* 111, and *one after another* 70; however, with one intervening word the figures change considerably, to: *one ___ after the other* 15; *one ___ after another*: 112. Relative examples from the corpus are:

‘He had smoked one cigarette after the other, holding them cupped in his palm...’

‘... by which time she had calmed down a little, but was smoking one cigarette after another.’

Not all phrases looked for in the BNC were found in alternative forms. Examples of hypothesized alternatives which were *not* found, are: *to talk (etc.) nineteen to a dozen*, *in a blink of an eye*, *to get a hang of sth*, and *to take the back seat*. Of course, absence from the BNC does not mean that a hypothetical alternative version is never used in the language or could not be found in another corpus.

2.2 Semantically Transparent Phrases

The specific phrases discussed in this section are examples of items noted by the author over a period of time, and representing different phraseological types. They are: the collocational verb phrase *put sth to a/the vote*, scalar phrases such as *to an/the inch* and *to a/the mile*, the numerical frame *in a/the ratio of [3:5]*, and the phrase *as a/the result of*.

With regard to the first of these phrases, the total number of corpus tokens (with either *a* or *the*) is 72, and *the* is considerably more frequent than *a* (57:15). These overall figures, however, hide an even greater difference, which relates to the spoken texts in the BNC. Actually, ‘spoken’ here means just certain types of language, since all tokens were found in the transcripts of meetings and discussions of various kinds. Here, the ratio of *the* to *a* increases to 36:2. In quite a number of instances the phrase is introducing an actual vote, which is about to take place or is being suggested (e.g. ‘Can I put that to the vote?’). It would seem, then, that *put to the vote* is very much the preferred form in these situations, though the figures may be inflated by repeated use of the phrase by individual chairpersons and similar. It should also be pointed out that the two tokens

with *a* are also used to refer to an actual vote: ‘So I think we should put that to a vote’; ‘I’ll put this to a vote then’.

In the written BNC texts, the ratio of *the* to *a* is 21:13. In the majority of cases, the two forms appear to be interchangeable, as in the following extracts:

‘... when a final version of the draft platform, retaining intact the proposals outlined by Gorbachev, was put to a vote by a show of hands on Feb. 7 it was adopted with only one vote against and one abstention.’

‘When the issue was put to the vote in the House of Commons we had a majority of 127 and the campaign against us was left in ruins.’

Occasionally there is modification of one sort or another, and *put to a vote* appears to be the only possibility (e.g. ‘to put the work of the government to a vote of no confidence’, ‘Otherwise, the issue will be put to a vote which also includes 20 smaller nations who lack Test status.’).

I turn now to the phrases *to an/the inch* and *to a/the mile*, which are mainly used in connection with the scale of maps. Examples of the former are:

‘With a scale of twenty-two miles to an inch, the escarpment should be massive...’

‘Western Europe is scaled at a useful 12 miles to the inch.’

Examples of the second phrase are:

‘Though I had some small scale maps at three inches to a mile for some areas...’

‘It is on what was for the period the large scale of almost one inch to the mile.’

The phrases *to an/the inch* were also found in a slightly different use, as in the following examples:

‘With synthetic thread, polyester for example, and rather long stitches at 4 to the inch or 6 mm each...’

‘... and that usually the eight points, which was meant to say there were eight threads to an inch.’

The next phrase is *in a/the ratio of*. Often, either article may be used, as in the following comparable contexts from the BNC:

‘Zero dividend preference shares are also being issued in a ratio of 37 for every 63 ordinary shares.’

‘... applicants for the Institute of Advanced Motorists test are in the ratio of one woman to four men.’

Where the word *ratio* is preceded by a classifying premodifier, *a* seems to be the only possibility, as in the following example: ‘The labelled Watson and unlabelled complementary strand were mixed in a concentration ratio of 1:2 respectively.’ In the case of other premodifiers, *the* is still an option, as in: ‘This difference is in the approximate ratio of 2:1.’

Lastly, it is to be noted with regard to *in a/the ratio of*, that the continuation of the phrase (or rather its completion) must be a numerical relationship; it cannot just include the names of the entities being compared, as in the following example, where only the

word *the* is possible: ‘... there is a remarkable constancy in the ratio of one element to the other’.

Turning now to the phrases *as a result of* and *as the result of*, these sometimes appear to be completely interchangeable, as can be seen by comparing the two extracts below.

‘If a radiator starts to leak (usually as the result of internal corrosion), try adding a radiator sealant at the feed-and-expansion tank.’

‘Where problems associated with water penetration have occurred on mastic asphalt roofs, it is usually as a result of failure from one or more of the following:...’

By far the more usual form of the phrase is *as a result of*, with 5149 tokens in the BNC as opposed to 338 for *as the result of*. These are raw frequency figures for the word strings, and about 20% of tokens of ‘as the result of’ represent other phenomena, especially the collocation *the result of* preceded by *as*, itself dependent on a previous verb (e.g. ‘There are many baffling viral diseases that cannot be readily explained as the result of an acute infection...’). By contrast, in a 10% sample of tokens of ‘as a result of’ just two items were irrelevant to the phrase *as a result of*.

3 The Frame *a(N)/the N₁ of n₂*

The lexico-grammatical frames *a(n) N₁ of n₂* and *the N₁ of n₂* sometimes overlap; that is, they can include the same *N₁* without a change in the resulting phrasal meaning or function. One such noun is the word *chance*, as is used in the following corpus extracts:

‘I think the best I’ve qualified has been 12th and I have never finished a race, so I am due something. I believe we have a chance of a good points finish, but of course it all depends on how we qualify on Saturday.’

‘But, if all goes according to plan, we have the chance of a title unification fight with even more appeal.’

The following pair of examples, again with the word *chance*, are even closer in their meaning and phraseology:

‘... has had the first tablets which may give her a chance of a normal life.’

‘He admitted that Estella was his housekeeper’s daughter, adopted by Miss Havisham to give her the chance of a better life.’

The word *chance* has a number of slightly different meanings, and in the above examples the sense is, quoting from [9], ‘a possibility of something happening, especially something that you want’. If we wish to rationalize the presence of either article in such phrases, the most obvious explanation is the fact that in the versions with *a chance of*, the word *a* is present because new information is being introduced, while in the versions with *the chance of*, the word *the* is being used cataphorically (though on a phrasal level, not textually). This phrasal phoric relationship is discussed by Willemse in some detail in [11], and I will now outline some of the points in this article that are of relevance to the present discussion. (For the use of the term ‘esphoric’ in Willemse’s article, see [12].)

Willemse's study revolves around one specific frame, which can be referred to as *the N₁ of a/an N₂* (e.g. 'the lights of a car'). The focus of the article is 'NPs involving a forward phoric relation', in which there are 'two discourse referents rather than only one'. Of specific relevance to the present paper, is the fact that in all 200 examples analyzed (from the Bank of English),

'... neither referent had been mentioned at an earlier point in the discourse before the esphoric NP₁, i.e. neither NP₁ nor NP₂ maintained any anaphoric relations with the preceding discourse context. This is of course not unexpected for NP₂, since it is an indefinite NP. However, it is remarkable for the definite NP₁...' (p. 329).

The relevance of this to the current study is the fact that it draws attention to the non-anaphoric nature of phrase-initial *the*. In phrases where this occurs, it would be reasonable to expect that the (often non-anaphoric) word *a* might also sometimes be found instead of *the*, thus allowing for both *the chance of a* and *a chance of a*.

This is not to say, of course, that the two versions will always be used indifferently in the language. As has been seen above in the case of fairly fixed phrases, there will often be restrictions and preferences for one or the other form. In the case of the word *chance*, and with reference just to phrases in which the second noun is preceded by *a/an* (as in Willemse's study), we find that in the BNC there are 27 tokens of *a chance of a* and 123 of *the chance of a*. Very often they appear to be interchangeable, though not always. For example, 'a chance of a' sometimes forms part of the longer phrase *in with a chance (of a...)*, while there is no evidence in the corpus of a parallel phrase 'in with the chance (of a...)'.

A number of first nouns can be found with either article in this frame; for the purposes of this paper I will refer to a few of the nouns mentioned by Willemse. In his paper he describes the various 'conceptual relations' which motivate forward bridging in his database. These are divided between the general categories of 'possessive relations' and 'contiguity relations'. One sub-category of the former is 'kinship relations', and a specific contextualized example is, 'A small urchin, pushing a bicycle, was with him. It was Wali Jan, from a mile down the valley; *the son of a smallholder*.' (Willemse's italics). Looking for the phrases *the son of a* and *a son of a* in the BNC, we find, not unexpectedly, that the former is by far the more frequent form. However, the important thing to note is that there are at least a few examples of *a son of a*. Examples of each phrase are:

'Mr Morton was born in South Africa, the son of a Scottish oil executive who had married into a local Afrikaner family.'

'Francis Bacon, well known for his capacity to drink nearly everyone under the table, left his friend John Edward, a son of a publican, £10 million in his will, which was published last month.'

With regard to 'contiguity relations', Willemse tells us that the most frequent type in his database is that of 'causal relations' (p. 348). One of the contextualized examples is, 'He was a tall, gaunt Scott in his mid-fifties with thinning hair and a pronounced limp on his left leg, *the result of an injury sustained during the Korean war*.' (Willemse's italics). In this particular example, 'a result' sounds less probable, and a search in the BNC suggests that the reason is its position in the sentence, grammatically speaking. The frequency figures for the word strings *the result of a* and *a result of a* are, respectively

473 and 350. However, if we make the same searches but with a preceding comma, as in the Willemse example, then we find that there are 28 tokens of ‘..., the result of a’, but none at all for ‘..., a result of a’.

In contrast to this, there are other situations in which the two phrases do seem to operate in parallel. The following are examples:

‘The occupation of the West Bank was a result of a war launched by third parties.’

‘Mrs Thatcher’s unwilling departure was the result of a combination of factors.’

Lastly in this section, I will give examples of two other words involving causal relations, *consequence* and *product*. Both are found in the BNC with each of the articles, and the following are examples:

‘The shortages were primarily a consequence of a blockade by Azerbaijan...’

‘This has been shown by lower prevalence of antibodies to toxoplasma in immigrants to Paris than in women of French origin, a factor suggested to be the consequence of a cultural preference for poorly cooked meat.’

‘As a gourmet you certainly think of yourself as a product of high civilization...’

‘Above all, worker radicalism was not the product of furious *démarches* by Bolshevik leaders against trade unionism...’

4 Concluding Remarks

The words *a/an* and *the* are usually thought of as being opposites, and their traditional labels, the indefinite and definite article, help to underline their very different functions. However, perceiving them in this way, and as part of a restricted grammatical paradigm, is to ignore their diverse functional and phraseological roles. Sinclair affirms that *the* and *a* each have ‘a word class all to itself’ ([3] p. 165), and I would fully endorse this opinion. In the same publication he writes that:

‘The very frequent words of English form a large proportion of any text, and yet their particular qualities are not fully recognized. They are not given adequate provision in theories of language, and their role is not very clearly described in either grammars or dictionaries, both of which take a somewhat partial view of their behaviour’. (p. 157)

It is in the context of this general research need that I am examining the phraseological overlap between *a* and *the*, which is one more piece in the puzzle of understanding the usage of the two words.

Exemplification of the phenomenon has been varied in this paper, but at the same time it has been limited. A few of the many phrases which have not been mentioned are *to the/a layman*, *(see) a/the need for STH*, and *(have) a/the right to (do) STH*, and only a few examples have been given of the nouns which make up the combined frame discussed in Sect. 3. Many phrases, and many contexts, would need to be examined in order to have a clear picture of the extent of the phenomenon, and also to understand why alternative articles are possible in some cases but not others.

Methodology-wise, it is to be noted that using corpus evidence to study very specific patterning of the words *a/an* and *the* is by no means straightforward. On the one hand,

there is sometimes insufficient data, as is the case with some idioms; on the other hand, where there is an abundance of material, it may be necessary to study many contextual environments, one by one, to discover whether uses with *a/an* and *the* are genuinely comparable. Asking computer software to recognize relatively simple patterns is by no means enough; for example, word strings such as ‘a model of’ and ‘the model of’ have a number of meanings. A further point is that analysis of corpus data involves a degree of personal judgement, since sometimes one has to assess whether a specific example would have been possible, with the same communicative effect, if the other article had been used; there is nothing unusual about this, however, since all corpus-based lexicology involves a degree of human intervention.

Finally, I would point out that it is relatively unusual in corpus-aided research to focus on what specific words have *in common*; we usually use the corpus to tease out the differences between words, phrases and structures, rather than the opposite. At the same time, however, focussing on the similarity between words, in this case the articles, can also show up differences in the phraseology of the words they combine with.

References

1. Berry, R.: Collins Cobuild English Guides: 3 Articles. HarperCollins Publishers, London (1993)
2. Moon, R.: Fixed Expressions and Idioms in English: A Corpus-Based Approach. Clarendon Press, Oxford (1998)
3. Sinclair, J.: A Way with Common Words. In: Hasselgård, H., Oksefjell, S. (eds.) *Out of Corpora: Studies in Honour of Stig Johansson*, pp. 157–179. Rodopi, Amsterdam (1999)
4. Sinclair, J.: The Search for Units of Meaning. *Textus Eng. Stud. Italy* **IX**, 75–100 (1996)
5. Collins COBUILD Dictionary of Idioms. HarperCollins Publishers, London (1995)
6. Cambridge International Dictionary of Idioms. Cambridge University Press, Cambridge (1998)
7. Hanks, P.: Dictionaries of Idioms and Phraseology in English. In: Corpas Pastor, G. (ed.) *Las Lenguas de Europa: Estudios de fraseología, fraseografía y traducción*, pp. 303–320. Granada, Editorial Comares (2000)
8. Moon, R.: Frequencies and Forms of Phrasal Lexemes in English. In: Cowie, A.P. (ed.) *Phraseology: Theory, Analysis, and Applications*, pp. 79–100. Clarendon Press, Oxford (1998)
9. Oxford Advanced Learner’s Dictionary. 9th edn. Oxford University Press, Oxford (2015)
10. Sinclair, J.: Collocation: A Progress Report. In: Steele, R., Threadgold, T. (eds.) *Language Topics: Essays in honour of Michael Halliday*, vol. II, pp. 319–331. John Benjamins, Amsterdam (1987)
11. Willemse, P.: Esphoric *the N of a(n) N*-nominals: Forward bridging to an indefinite reference point. *Folia Linguistica* **40**(3–4), 319–364 (2006)
12. Martin, J.R.: *English text: System and structure*. John Benjamins, Amsterdam (1992)