Time-measurement constructions in English: a corpus-based exploration

Melanie J. Bell & Carmen Portero Muñoz

Anglia Ruskin University / University of Córdoba

Abstract

Time-measurement expressions such as *five-year plan, 10 years' time* and *25 years service* occur frequently in English. All such expressions consist of a cardinal numeral, followed by a time-noun (N1) then a second noun (N2). The time-noun has one of three orthographic forms: the bare-form, the S-form with apostrophe or the S-form without apostrophe. Using a dataset of 17591 time-measurement tokens from the *British National Corpus* and mixed-effects logistic regression modelling, this chapter tests the hypothesis that these three orthographic forms represent three different constructions. Our first model, using only expressions with S-form N1, shows that the presence or absence of an apostrophe is not correlated with any other formal or semantic property that would justify the recognition of two constructions. In contrast, our second model using the whole dataset, shows that bare-form N1 and S-form N1 (with or without apostrophe) are highly correlated with aspects of both form and meaning. In our dataset, 96% of tokens with bare-form N1 have a countable N2 and 87% also follow a determiner. Conversely, 94% of tokens with S-form N1 have an uncountable N2, and 91% also lack a determiner. We conclude that these clusters of properties represent distinct pairings of form and meaning, and are therefore characteristic of two different constructions, which we call the TIME-MEASUREMENT COMPOUND construction and the TIME-MEASUREMENT construction respectively. The TIME-MEASUREMENT COMPOUND construction (*five-year plan*) has the distribution of a nominal; semantically, it denotes a kind of bounded entity (N2) with some relation to numeral-N1, usually duration. The TIME-MEASUREMENT construction (*10 years' time*, *25 years service*) has the distribution of a noun phrase; semantically, it denotes a

quantity (numeral-N1) of some unbounded entity (N2). The chapter ends with a qualitative exploration of the central and more peripheral representatives of the two constructions, including borderline cases.

1. Introduction[1]

The aim of this paper is to propose the existence of two time-measurement constructions in English. Our analysis is based on the types of expressions exemplified in (1), where each example is taken from the *British National Corpus* (BNC):

(1) a. *10 days free trial* (BNC CFT 33)

11 year reign (BNC CJC 450)

12 years operation (BNC H0B 563)

24-hours every-day operation (BNC HBD 300)

one-day general strike (BNC HKX 861)

one-month's delay (BNC HNL 443)

ten second intervals (BNC CGD 1695)

[1] **Acknowledgements:** We are very grateful to everyone who assisted in the preparation of this chapter. Sebastian Hoffmann provided corpus frequencies, Martin Schäfer helped with R code and semantic terminology, and John Potter and Satu Vartiainen checked the parsing of the data. We also thank the two anonymous reviewers who provided such thoughtful and constructive feedback, and above all the editors of this volume for their humbling patience. The work was partly funded by a European Science Foundation NetWords travel grant to Carmen Portero Muñoz, who conducted the work within the framework of the research group HUM693 Lingüística Cognitiva y Funcional (LINCOFU) (Autonomous Government of Andalusia).

three weeks' rehabilitation (BNC CBG 11628)

b. *14-year salary* (BNC A9V 194)

forty four hours wages (BNC GYU 198)

one weeks papers (BNC J1H 3501)

six months' interest (BNC G2K 1179)

In form, all such expressions consist of a cardinal numeral, followed by a time-noun (N1), then a second noun (N2), with or without an intervening modifier or modifiers. Semantically, the combination of the numeral plus N1 (num-N1) represents a measurable quantity of N2. Often, this quantity is a duration, as in (1a), but it can also indicate a value or other quantity corresponding to an amount of time, as in (1b). However, despite their shared characteristics, these expressions vary in the orthographic form of N1. In some cases, N1 is written with both a final *s* and an apostrophe (*one-month's delay, six months' interest*), in other cases, N1 is written with final *s* but no apostrophe (*12 years operation, one weeks papers*), while in a third type, N1 is written with the bare-form, i.e. neither *s* nor apostrophe (*ten second intervals, 14-year salary*).

The formal variation in time-measurement expressions has been noted in all the main reference grammars of English. For example, Quirk et al. (1985: 1333) point out the similarity between the examples shown in (2), which they analyse as having the forms indicated in square brackets:

(2) a. *a ten day absence* [singular]

b. *a ten-day absence* [hyphen + singular]

c. *a ten days absence* [plural]

d. *a ten days' absence* [genitive plural]

Payne and Huddleston (2002: 470) also note this similarity. In their analysis, expressions such as *two-hour* in *two-hour delay* are classed as compound adjectives and represent an alternative to the 'genitive', such as *an hour's delay*, as a means of expressing measure. However, there is little discussion in either text, or elsewhere in the literature, of the factors that lead to one form being chosen rather than another. The present chapter fills this gap. Specifically, we answer the following research question: Are the different orthographic forms of the time-noun in English time-measurement expressions indicative of different constructions?

We use mixed effects logistic regression modelling of a large sample of data from the BNC, to analyse the formal, semantic, and distributional properties of the expressions exemplified in (1). Our first model, using only expressions with S-form N1, shows that the presence or absence of an apostrophe is not correlated with any other formal or semantic property that would justify the recognition of two constructions. In contrast, our second model, which includes all three orthographic forms, shows that the choice

between bare-form N1 and S-form N1 (with or without apostrophe) is highly correlated with aspects of both form and meaning. We therefore conclude that there are two main types of English time-measurement expression, representing two different constructions: a TIME-MEASUREMENT construction, whose typical exemplars have S-form N1, with or without apostrophe, and a TIME-MEASUREMENT COMPOUND construction, whose typical exemplars have bare-form N1. Finally, we support this analysis with a more qualitative exploration of the data, showing that the constructions we propose satisfy the widely adopted criteria for constructional status proposed by Goldberg (2006).

The rest of the chapter is organised as follows. In Section 2, we review what has been written about time-measurement expressions and similar constructions, concluding that there is currently no adequate analysis. In Section 3, we describe how we assembled and coded a dataset of pertinent examples from the BNC to enable us to fill this gap. In Sections 4 and 5, we report two statistical analyses, using our dataset to test whether variations in the form of N1 are associated with variation in meaning such that we can identify distinct form-meaning pairings. In Section 6, we present a qualitative exploration of the central and more peripheral representatives of the two constructions, including borderline cases. Finally, Section 7 is a summary of our conclusions.

2. Background

2.1 Construction Grammar

Our theoretical assumptions are based on the tenets of constructionist linguistics summarised by e.g. Croft and Cruse (2004), Goldberg (2013). We conceptualise language as a network of constructions, i.e. pairings of form and function, in which there is no modular separation of morpho-syntax from lexicon, and surface form is associated directly with meaning. We further theorise that language is usage-based (cf. Bybee 2013; Diessel 2019), so that constructions emerge in the minds of speakers as generalisations over the sum of the individual's linguistic experiences. Our definition of CONSTRUCTION is that of Goldberg (2006):

> Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency.
> Goldberg (2006: 5)

Although, within a usage-based theory of language, it is logical to assume that speakers will store frequent patterns even if they are fully predictable, the frequency criterion in Goldberg's (2006) definition of constructions is problematic; as Traugott and Trousdale (2013: 11) point out, "sufficient frequency" is difficult to operationalise. We will take the view that if a combination of form and function occurs in a representative sample of a language with sufficient frequency for the pairing to emerge as significant in a statistical analysis of that sample, then speakers of the language might also be sensitive to the pairing. In other words, we assume that the relative strengths of statistical patterns in the language reflect the likelihood of those patterns being entrenched as constructions in the minds of speakers. Regarding the extent to which aspects of linguistic form or function are predictable from an expression's own structure or other established constructions, we find it useful to apply the strategies listed by Hilpert (2019: 14-22). Namely, a construction can be recognised as non-predictable if it deviates from canonical patterns, carries non-compositional meaning, has idiosyncratic constraints, or has collocational preferences.

By 'function', we mean the conceptual content of an utterance in context, including its semantic and pragmatic meanings as well as its syntactic role. Our notion of 'construction' therefore takes into account not only the internal form and meaning of a linguistic expression, but also its syntagmatic and paradigmatic distribution in relation to other constructions

(cf. Goldberg 2006: 5). Nevertheless, we agree with Croft (2001: 13) that there is a risk of circularity in distributional analysis if categories such as word classes are defined in terms of constructions which are conversely defined in terms of the categories they contain. To avoid such circularity, it is necessary to define substitution classes in terms of specific combinations of form and function. For example, one such substitution class, which plays a central role in the expressions discussed in this chapter, is the English NOUN construction. We define the NOUN construction as the minimal form that can combine with [ðə] to refer to an entity, and the NOMINAL construction as any longer string that can occupy the same slot (cf. Bell 2012).

Because our research question concerns differences in orthographic representation, the notion of constructional form requires more discussion. By 'form', we mean the phonological form of an expression, which we assume arises as a generalisation over the phonetic forms of individual utterances experienced by a speaker. When we refer to the S-form of a noun, we are therefore referring to phonological forms ending in one of the allomorphs /s/, /z/, or /ɪz/. A child acquiring a language first learns to associate meaning with sounds, or in some cases gestures, and the orthographic form only comes later, when the child learns to read and write. In a usage-based view of language, it is inevitable that spelling will have a back-wash effect on the mental grammars of literate speakers. However, we

take the view that, unless variation in spelling corresponds to variation in some other aspect of form or function, it may reflect the prescriptive pressures of formal education as much as the writer's mental grammar. The orthographic variation in the examples in (1), namely the use of digits or words to express the numeral, the presence or absence of an apostrophe, and the presence or absence of a hyphen, does not affect the phonological form of these expressions. Furthermore, numerals, hyphens, and apostrophes all feature prominently in style guides for written English, such as the *University of Oxford Style Guide*, suggesting that writers may refer to external rules as much as their internal representations when using these features. We therefore do not regard these orthographic differences *per se* as sufficient basis for the recognition of different construction forms.

There are many cases in English where orthographic variation between homophones, i.e. orthographic variation in the absence of phonological variation, is clearly correlated with a difference in meaning. For example, in the language generally, the presence or absence of an apostrophe marks the semantic difference between a common case plural, meaning 'more than one', and a genitive plural, broadly meaning 'possessed by more than one'. This difference is also correlated with differences in distribution; for example, in *the cats are in the kitchen*, *the cats* functions as the subject of the clause and is followed by a verb, whereas in *the cats' food is in their bowls*, *the cats'* functions as a determiner and is followed by a

noun. In other words, although *cats* and *cats'* are homophonous, a learner of the language could discern a difference in meaning on the basis of the difference in distribution, and the presence or lack of an apostrophe reflects this difference. In other cases, purely orthographic differences appear not to be indicative of differences in meaning, but rather to reflect conventions of style. For example, style manuals advise that numerals from one to ten should be written as words whereas numerals above ten should be written as digits unless they occur sentence initially. However, the choice between these forms does not reflect any difference in meaning: *ten years* means the same as *10 years*. Our first question regarding time-measurement expressions is whether the presence or absence of an apostrophe in these expressions reflects a difference in meaning, as it does in the common case plural and genitive plural generally, or whether it is simply an orthographic convention similar to the difference between numerals written as words and numerals written as digits.

There is some evidence that common case plural /S/ and genitive plural /S/ are not in fact true homophones but differ systematically in duration and are therefore formally distinct at the level of phonetics (Plag et al. 2020). However, it is unclear whether this difference occurs in normal spontaneous speech, since Plag et al.'s (2020) data was taken from an experiment in which sentences were read aloud, meaning that participants would have been subject to the visual cues provided by the presence or

absence of an apostrophe. And even if the difference does occur in spontaneous speech, it is not known whether the 7 or 8 millisecond difference in duration is long enough for listeners to be sensitive to it. Furthermore, it is an open question whether a durational difference would emerge between our time-measurement expressions with and without an apostrophe; in Plag et al.'s (2020) data, apostrophe use was accompanied by other distributional cues which would not occur in the time-measurement context. For all these reasons, we do not assume that the presence or absence of an apostrophe in time-measurement expressions is correlated with any difference in the acoustic signal; hence, unless the orthographic variation can be shown to correlate with some other aspect(s) of form and/or function, we will not regard variation in apostrophe use as indicating two different constructions.

As far as we know, the time-measurement expressions exemplified in (1) have not previously been analysed in any detail, either in the framework of Construction Grammar, or elsewhere in the literature. However, they are mentioned briefly in all the main reference grammars of English, and the Constructionist literature includes descriptions of several other constructions that are similar in terms of form or meaning or both. In the rest of this section, we will discuss four such constructions, comparing their form and meaning to our time-measurement expressions, and relating them to the descriptions given in the reference grammars. We base our

representations of these constructions on the "fairly informal description of the form and meaning parts" of Hoffmann and Trousdale (2013: 1), using square brackets to indicate constituency, and subscript tags to indicate the more general construction to which a constituent belongs.

2.2 The S-GENITIVE construction

The first construction that resembles our time-measurement expressions is the S-GENITIVE construction shown in (3).

(3) S-genitive construction (Hilpert 2019: 61)
e.g. *John's book*, *the country's president*
$[[X\ S]_{NP}\ N]_{NP}$ – 'X possesses N'

In terms of orthography, this construction matches time-measurement expressions where N1 is written with *'s* or *s'*, e.g. *one-month's delay*, *six months' interest*. In terms of phonology, it also matches those where N1 is written with *s* without apostrophe, e.g. *10 days free trial*, *one weeks papers*. The apostrophe is sometimes included in representations of the S-GENITIVE construction (e.g. Gries & Stefanowitsch 2004), but since we are taking the salient form to be the phonological form, we have not included it in (3).

It is mainly in their sections on the genitive that the reference grammars mention time-measurement expressions, and specifically in the context of the so-called 'genitive of measure', exemplified in (4). Biber et al. (1999: 296) note that this construction frequently occurs with time-nouns to express duration (4a and 4b), distance (4c), amount (4d), or value (4e).

(4) a. *an hour's delay*

b. *several weeks' vacation*

c. *12 hours' journey*

d. *ten years' supply*

e. *half an hour's income*

Quirk et al. (1985: 1333, 1276) analyse 'measure genitives' such as *ten days'* in *a ten days' absence* as attributive modifiers, in contrast to most other genitives, which they regard as determinatives. Payne and Huddleston (2002: 470) share this analysis. They acknowledge that, when the time expression occurs initially in the noun phrase, as in *an hour's delay* or *one week's holiday*, it resembles a determiner. Nevertheless, the acceptability of the examples in (5), where the time expression *follows* a determiner, leads them to conclude that all 'measure genitives' are in fact modifiers.

(5) a. *this* [*hour's delay*]

b. *a second* [*one hour's delay*]

This conclusion rests on two assumptions. Firstly, it assumes that any determiner preceding the time expression belongs to N2 rather than N1. However, this is not always easy to establish. In *this hour's delay*, for example, the demonstrative *this* could also be taken to refer to the noun *hour*, with the bracketing shown in (6):

(6) [*this hour*]*'s delay*

Since the scope of the determiner is not constrained by number agreement, it is hard to know how one could be sure of the intended bracketing in such a case, except perhaps in the most marked or contrastive of contexts. Secondly, even if we decide that the demonstrative (*this*) belongs with N2 (*delay*), the conclusion that N1 (*hour's*) must therefore be a modifier rests on the assumption that N2 cannot take both the demonstrative and N1 as determiners. This assumption follows from an analysis of N1 as genitive, since *s*-genitive determiners occupy the same central determiner slot as e.g. articles and demonstratives, and there can only be one central determiner for any noun (Quirk et al. 1985: 254, 326; Biber et al. 1999: 258ff.). However, in an alternative analysis where N1 is taken to be not a genitive but a

quantifier, it would be possible to analyse both the demonstrative and N1 as determiners of N2. On this view, N1 would be a quantifying post-determiner, analogous to e.g. *few* in *these few days* (Quirk et al. 1985: 262-4; Biber et al. 1999: 258ff.).

Regarding factors that might correlate with the form of N1, an observation by Payne and Huddleston (2002: 470) suggests to us that the countability of the second noun might be relevant, in conjunction with the presence or absence of other phrasal constituents. They state that 'measure genitives' (i.e. types with S-form N1) cannot occur initially in the noun phrase when the second noun requires a count interpretation. To illustrate this, they claim that **an hour's game of squash* would be unacceptable, the preferred form being *a one-hour game of squash*. However, this seems to contradict the examples given by Quirk et al. (1985: 1333), e.g. *a ten days absence, a ten days' absence*, where the indefinite article does not agree with either the cardinal numeral or the plural N1 and must therefore have scope over N2, *absence*, which thus has a count reading. We note, however, that whereas Quirk et al.'s (1985: 1333) examples all include a numeral, some of Payne and Huddleston's (2002: 470) do not. Overall, there is no clear account in the previous literature as to whether variation in the form of N1 is related to the countability of N2, or what other factors, such as the presence or absence of a numeral, determine the form produced in any given case. Again, if the different forms of N1 are correlated with differences in

meaning or function then we would conclude that they represent different constructions.

The status of the apostrophe in time-measurement expressions is mentioned almost in passing by both Quirk et al. (1985: 325, note [b]) and Biber et al. (1999: 293). Both sources state that when the temporal noun is plural, as in *several weeks' vacation*, the apostrophe is sometimes "omitted", e.g. *several weeks vacation, ten years imprisonment*. Biber et al. (1999: 293) suggest that the variation is "best regarded as involving a choice between genitive and common case […] rather than between alternative spellings of the genitive". In their analysis, the form of the time word in expressions like *several weeks vacation* is common case plural. They further suggest that the choice between the two forms might depend on the lemma of N2 and the number of N1. For example, they report that, in the *Longman Spoken and Written English Corpus*, the apostrophe tends to be used regularly when the second noun is *time* but omitted with *imprisonment*. However, when the time word is singular, as in *an hour's delay*, they find that the form with an apostrophe is regularly used (Biber et al. 1999: 293). If the different orthographic forms reflect systematic differences in meaning, for example in terms of the number of N1 or the lemma of N2 (cf. Biber 1999: 293), then there are grounds to recognise two constructions, one written with the plain S-form and one written with *s* plus apostrophe. On the other hand, if no such correlation between orthographic form and meaning

can be found then we would regard these expressions as belonging to a single construction with two orthographic variants, possibly related to stylistic factors.

We now turn to the overall meaning of the S-GENITIVE construction and specifically the 'genitive of measure'. Quirk et al. (1985: 322) and Payne and Huddleston (2002: 470) paraphrase *a ten days' absence* and *an hour's delay* with the *of*-constructions in (7a) and (7b) respectively.

(7) a. *an absence of ten days*

b. *delay of an hour*

This interpretation follows straightforwardly from an analysis of N1 as an *s*-genitive, by analogy with e.g. *the bride's mother ~ the mother of the bride*. In contrast, Bauer et al. (2013: 143) consider cases like *an hour's delay* to be semantically partitive. Although they include this expression as an example of an *s*-genitive, they liken its meaning to *a lump of cheese*, as shown in (8). On this view, the meaning of *an hour's delay* is not 'delay of an hour' but rather 'an hour of delay'.

(8) a. *an hour's delay*

b. *a lump of cheese*

However, if Bauer et al.'s (2013) interpretation is correct, then it is difficult to maintain a genitive analysis of the time word because the inflectional mark appears to be added to the 'wrong' noun. By analogy with e.g. *gravity of Earth ~ Earth's gravity*, we might expect to find *hour of delay* alternating with **delay's hour* rather than *hour's delay*.

In a Construction Grammar approach the problem of deciding which prepositional expression alternates with the *s*-genitive does not arise, since each constructional form is regarded as having its own meaning. There is in fact some empirical evidence (Gries & Stefanowitsch 2004) that the S-GENITIVE and genitive *of*-constructions are semantically distinct (cf. also Langacker 1995; Stefanowitsch 1998, 2003). It is well documented that genitives can have a variety of interpretations; for example, *John's book* could represent a book that John owns, a book that he has on loan from a library, or a book that he has written, amongst other possibilities. However, Taylor (1989) argues convincingly that the central meaning of the genitive is one of possession, with the other interpretations being sense extensions of this central meaning. For example, if John has written a book, then he possesses the authorship and possibly copyright. Stefanowitsch (2003) shows further that, while the *s*-genitive basically encodes possession, including familial and social relations, the *of*-construction encodes classificatory and part-whole relations. In the case of our time-measurement expressions, despite the formal resemblance to the S-GENITIVE construction,

it is difficult to see how the meaning of 'possession' would apply. Even if the concept of possession is extended to incorporate the sense of occupation, then in *one month's delay*, for example, the delay seems to us to occupy one month, rather than the month occupying the delay. Overall, we conclude that the S-GENITIVE construction resembles some of our time-measurement expressions in form but not meaning.

2.3 The MEASURE NOUN PSEUDO-PARTITIVE construction

Expressions like *a lump of cheese* are more accurately classed as pseudo-partitive, rather than partitive. Whereas true partitives refer to part of a bounded entity, e.g. *a lump of that cheese*, *a year of their marriage*, pseudo-partitives refer to an amount of some unbounded entity, e.g. *a lump of cheese, a year of marriage*. Pseudo-partitives related to time-measurement, e.g. *an hour of delay*, belong to the sub-type classed by Keizer (2007: 109, following e.g. Vos 1999 for Dutch) as 'measure-noun' constructions (9):

(9) Measure noun pseudo-partitive construction (Keizer 2007: 109ff.)
e.g. *a pint of beer, twenty-nine years of marriage*
$[[\text{Det N1}_{\text{measure}}]_{\text{NP}}$ *of* $\text{N2}]_{\text{NP}}$ – 'quantity Det N1 of N2'

In contrast to the S-GENITIVE construction, the MEASURE NOUN PSEUDO-PARTITIVE construction resembles our time-measurement expressions in meaning but not in form. Formally, the PSEUDO-PARTITIVE construction always includes the string *of* between N1 and N2, which is absent from time-measurement expressions. However, semantically, there is a clear similarity between e.g. *29 years of marriage* and *three weeks' rehabilitation*.

2.4 The PHRASAL COMPOUND construction

As well as resembling the *s*-genitive, time-measurement expressions with S-form N1 also resemble the PHRASAL COMPOUND construction shown in (10).

(10) Phrasal compound construction (Hilpert 2019: 80)
e.g. *over the counter drugs*
$[[X]_{XP}\ N]_N$ – 'a kind of N with some relation to X'

Orthographically, the PHRASAL COMPOUND construction matches time-measurement expressions written with the S-form without apostrophe, as in *forty four hours wages*, since *forty four hours* is a noun phrase, just as *over the counter* is a prepositional phrase. Phonologically, the construction also matches time-measurement expressions where S-form N1 is written with an

apostrophe, e.g. *six months' interest, three weeks' rehabilitation*. Types with bare-form N1 only resemble phrasal compounds if the numeral is 'one', e.g. *one-day general strike*. In terms of semantics, time-measurement expressions conform to the PHRASAL COMPOUND template in cases where the modifying phrase serves to mark out a particular kind of the entities represented by N2. For example, a 'one-day general strike' might be regarded as a kind of general strike that lasts one day, and therefore *one-day general strike* could be classed as a phrasal compound. In contrast, 'three weeks' rehabilitation' is not a kind of rehabilitation, but rather an amount, and *three weeks' rehabilitation* is not a phrasal compound.

The meaning of the PHRASAL COMPOUND construction is reflected in its syntactic distribution: compounds are nominals rather than noun phrases. This means that if a time-measurement expression is a phrasal compound, the string from numeral to N2 inclusive will constitute a nominal. And since nominals do not include determiners, any determiner preceding the numeral will be external to the time-measurement expression, and therefore belong functionally with N2. For example, consider the time-measurement expression *24-hours every-day operation* in the context shown in (11), where *operation* is being used in a countable sense.

(11) a. *Through a **24-hours every-day operation**, about 150 tons of flint are processed each week.* (BNC HBD 300)

b. [*a* [[*24-hours*]$_{NP}$ [*every day*]$_{NP}$ *operation*]$_{N}$]$_{NP}$

‘a kind of operation lasting 24 hours every day’

In (11a), the determiner (*a*) agrees in number with N2 (*operation*), so the noun phrase headed by *operation*, including the determiner, has the structure shown in (11b). In other words, in the context of (11), *24-hours every-day operation* is a nominal and has the form and function of a phrasal compound. Contrast this with the context shown in (12), where *operation* is being used in an uncountable sense.

(12) a. *This makes a grand total of 1,031 tons 14 cwts. for those first **12 years operation**.* (BNC H0B 563)

b. [[*those first **12 years***]$_{NP}$ *operation*]$_{NP}$

‘the initial 12 year period of operation’

In (12a), the determiner (*those*) agrees in number with N1 (*years*), so the noun phrase headed by *operation* has the structure shown in (12b). Thus, in the context of (12), *those first 12 years* is a quantifying noun phrase, and *12 years operation* is therefore not a phrasal compound; rather, it forms part of an expression with the phonological form, though not the meaning, of an *s*-genitive.

2.5 The MEASUREMENT AS MODIFIER construction

When N1 has the bare-form, as in *a ten day absence,* time-measurement expressions resemble the MEASUREMENT AS MODIFIER construction shown in (13).

(13) Measurement as modifier construction (Hilpert 2019: 15)

e.g. *a sixteen-year-old boy*

[(Det) [[Num $N1_{time}$ Adj] N2]$_N$]$_{NP}$

'N2 has magnitude Num N1 Adj'

In surface form, the MEASUREMENT AS MODIFIER construction resembles time-measurement expressions like *(a) one-day general strike*, where an adjective intervenes between N1 and N2. The difference is that in the MEASUREMENT AS MODIFIER construction, the adjective forms a constituent with num-N1, e.g. *sixteen-year-old*, whereas in expressions like *one-day general strike*, the adjective modifies N2 independently of num-N1. In terms of meaning, there is a clear parallel in that *sixteen-year* represents the age of the boy just as *one-day* represents the duration of the strike. However, whereas the attribute being measured is overtly expressed in the MEASUREMENT AS MODIFIER construction (e.g. age in the case of *sixteen-*

year-old boy), in a time-measurement expression, it is understood (e.g. duration in the case of *one-day general strike*). Thus, the MEASUREMENT AS MODIFIER construction resembles some of our expressions in both form and meaning but is not identical to them in either respect.

2.6 Interim summary

In summary, apart from a subset that are captured by the PHRASAL COMPOUND construction, the time-measurement expressions exemplified in (1) are not adequately described by any currently established construction of English. The reference grammars also reveal some uncertainty, or at least ambiguity, about their status. Overall, the puzzle that these various forms represent is encapsulated in the words of Rosenbach (2006: 114):

> Like descriptive genitives, measure genitives appear to be very *multifaceted constructions*. Future research will have to look in more detail at the structural properties of measure genitives, how they evolved historically, how precisely they tie into the category of *s*-genitives, and how they overlap constructionally with N+N sequences. (emphasis added)

In the rest of this chapter, we will partially[2] solve this puzzle by analysing a large number of exemplars of the relevant types and exploring how their formal and semantic properties cluster together. In particular, we will investigate what other semantic and structural features are associated with the different orthographic and phonological forms of N1, with a view to establishing whether these different forms represent different constructions.

3. Creating a database

3.1 Selecting examples

In order to model the factors that predict the form of N1 in time-measurement expressions, we needed a representative sample of such expressions. To this end, we used the Simple Query Syntax of the Lancaster Interface to the BNC (Hoffmann et al. 2008), to extract all sentences containing a string of the form shown schematically in (14).

(14) (quantifying expression)* cardinal-numeral time-noun (adverb)? (adjective)* (noun)+

[2] We only consider the synchronic form and function of time-measurement expressions and not 'measure genitives' more widely, nor their historical evolution.

In this template, *?* means ‘zero or one’, * means ‘zero or more’, and + means ‘one or more’. In other words, we extracted sentences containing strings in which an optional quantifying expression is followed by a numeral, then a time-noun, then an optional adverb, zero or more adjectives, and at least one more noun.

What we are calling ‘quantifying expressions’ are modifiers of the numeral, and include those classes of words called “downtoners” by Quirk et al. (1985: 597). These are ‘approximators’ such as *almost*, ‘compromisers’ such as *more or less*, ‘diminishers’ such as *only*, and ‘minimisers’ such as *barely*. The full set of quantifying expressions included in our corpus queries is shown in the first column of Table 1. Including these optional terms in our query syntax facilitated automatic parsing of the output in cases where they occurred. Likewise, including one or more final nouns in the query syntax facilitated identification of the head in nominal compounds, since it was always the final word in the hit. The ‘numeral’ in our queries was any sequence of letters or digits tagged as a cardinal numeral (e.g. *one, 3, fifty-five, 3609, 2.5*). We also included coordinated numerals with the forms ‘between x and y’, e.g. *between four and five*, ‘x to y’, e.g. *four to five* or *4-5*, and ‘x or y’, e.g. *‘four or five*’. The set of time-nouns included in our queries is shown in the second column of Table 1. We searched for both the bare-form and the S-form of these nouns, with or

without an apostrophe, and with or without a hyphen between the numeral and the time-noun. Apart from the quantifying expressions and time words, the corpus queries only specified parts of speech.

Table 1: Lexical items specified in the corpus queries

Quantifying expressions	**Time nouns**
barely, less than, more than, longer than, greater than, above, over, as much as, around, about, nearly, roughly, approximately, only, just, at least, at most, up to, the last, maximum, minimum, a maximum of, a minimum of, an average of, the equivalent of, a total of, another, a further, some, as little as, as few as, only just, shorter than, more or less	*year, month, week, day, night, hour, afternoon, minute, second*

The raw corpus hits were checked manually by the authors to exclude any in which the numeral plus time-noun did not express a quantity of the final noun. This left a total of 17,591 sentences. Some examples are shown in (15), together with the name of the BNC file in which they occur. The query hit in each sentence is highlighted in bold. Overall, the hit varied from three (15a-f; 15o) to nine (15n) words in length.

(15) a. *Never before had England beaten West Indies in two consecutive **one-day games**, let alone three ...*
(BNC ABR 1124)

b. *To celebrate the chorus's **25 year association** with the Festival ..., they perform three of the most ambitious works in the repertoire ...* (BNC EC4 394)

c. *... the acid output during each **15 minute period** was measured by titration to pH 7 ...* (BNC HU3 4204)

d. *... even that brisk extra **thirty minutes' walk** every day, will gradually show satisfying results.* (BNC B3G 1350)

e. *At Georgetown, Toyota keeps **three days' worth** of stocks of imported parts ...* (BNC ABE 2128)

f. *That meant she had a good **eight hours' start** before anyone need even think about her absence.* (BNC FNT 13)

g. *Students who undertake the **four-year sandwich course** spend the third year in industrial placement.*
(BNC B3C 1798)

h. *An optional **10-minute speed test** (2105) may be taken by candidates entering for the proficiency examination, without additional fee.* (BNC HBP 2000)

i. *It is worth reflecting what a most remarkable contribution women have made to Save The Children throughout its* ***seventy four years history****.* (BNC JNG 258)

j. *The consultant, Dr. Nigel Cox, has been a given a* ***one-year suspended jail sentence*** *for attempting to murder a terminally ill patient.* (BNC K21 1072)

k. *... four policemen in Guatemala City were convicted of the murder of a 13-year-old street child and sentenced to* ***between 10 and 15 years' imprisonment****.* (BNC A03 97)

l. *After no* ***more than five or ten seconds' pause*** *the door shivered under a hard kick ...* (BNC G03 2447)

m. *The role has evolved as a basic but flexible care worker, with* ***a minimum of three months skill-based training****.* (BNC HXT 471)

n. *In addition, employees may be granted* ***up to a maximum of three days' special leave****.* (BNC CHS 735)

o. *... all worked by men whose* ***seven-year apprenticeships*** *made them masters of their craft.* (BNC CHP 70)

The sentences were coded for a number of metatextual, orthographic, morphosyntactic, length, frequency, and semantic variables, as detailed in

the rest of Section 3. The full coded dataset is available at https://doi.org/10.25411/aru.14727663.

3.2 Metatextual categories

The following three variables are part of the metatextual mark-up of the BNC and were extracted along with the sentences. We included these variables in our modelling to control for possible stylistic effects, especially related to orthographic variation.

- DERIVED TEXT TYPE: the genre of text in which the sentence occurred. The possible values are 'academic prose' (WA), 'fiction and verse' (WF), 'newspapers' (WN), 'non-academic prose and biography' (WP), 'other published written material' (WO), 'unpublished written material' (WU), 'spoken conversation' (SC), and 'other spoken material' (SO)
- TYPE OF AUTHOR: possible values are 'sole', 'multiple', and 'corporate'. In the BNC, authorship of a written text is characterised as 'corporate' if it was produced by an organisation and the author(s) are not named. Authorship is characterised as 'multiple' if there is more than one named author.

- SEX OF AUTHOR: possible values are ‘female’, ‘male’, ‘mixed’, and ‘unknown’.

3.3 Orthography

Each sentence was coded for three binary orthographic variables; we could do this automatically on the basis of our search terms (e.g. whether we had included a hyphen in the syntax). To exemplify this coding, we will list the sentences from (15) that have a particular value for each variable. The sentences not listed have the other value in each case.

- APOSTROPHE: the presence (15d, e, f, k, l, n) or absence of an apostrophe following the time word
- HYPHEN: the presence (15a, g, h, j, o) or absence of a hyphen between the numeral and time word
- NUMERAL FORM: whether the numeral is expressed in digits (15b, c, h, k) or words.

3.4 Morphosyntax

The various constituents of the corpus hits and of the phrases containing them, were first extracted using the part-of-speech tags provided in the BNC

together with possible phrase structures. To identify premodifiers and quantifiers not in our original list, we searched leftwards from the hit until reaching a tag for a preposition, verb, or punctuation, any of which we expected to signal that we had passed the start of the noun phrase containing the hit. This search, and the parsing of the hit itself, was first carried out automatically using R (R Development Core Team, 2019). We also added the lemmas of N1 and N2 to the dataset. To do this, we used a frequency list provided by Sebastian Hoffman (Hoffmann et al. 2008) that included every wordform in the corpus with each of its part of speech tags plus the corresponding lemma. After automatic parsing and addition of lemmas, all 17,591 lines of data were checked and, if necessary, manually corrected by one of two student assistants who were studying English linguistics. As explained for the orthographic coding, examples from (15) are given for just one value of each binary variable, with the second value applying in the unlisted cases:

- N2 LEMMA: the lemma of the final noun, i.e. the singular form. In (15a-o) the lemmas are *game, association, period, walk, worth, start, course, test, history, sentence, imprisonment, pause, training, leave*, and *apprenticeship* respectively.
- N1 FORM: the form of the time word, either 'bare-form' (15a, b, c, g, h, j, o) or 'S-form'. Note that the coding 'S-form' includes those

written with an apostrophe as well as those written without an apostrophe.

- PREPOSITION: the presence (15a, c, i, k, l, m) or absence of an initial preposition in the phrase containing the hit. In other words, this variable codes whether the hit occurs in a prepositional phrase or a noun phrase.
- PREPOSITION TYPE: A distinction was made between prepositions of time (15c, i, l) and other prepositions (15a, k, m). Because the same preposition can have multiple uses, this variable was coded manually, taking the context into account. For example, in the context of … *the holding of elections* ***in one or two years' time*** (BNC K5M 12305), the preposition *in* means 'after a specified length of time', and preposition type was therefore coded as 'time'. In contrast, in the context of ... *a further cut* ***in 1–2 year fixed-rate deals*** (BNC CEL 1271), the same preposition collocates with the preceding noun to indicate a reduction (*a cut in sth*), and preposition type was therefore coded as 'other'.
- PREMODIFIER: the presence (15a, d, f, h) or absence of an adjective before the numeral but within the same phrase.
- DETERMINER: the presence or absence (15e, k, l, m, n) of a determiner before the numeral but within the same phrase.

- DETERMINER TYPE: the type of determiner, if any. The following values were coded: ‘numeral’ (15a), ‘indefinite article’ (15f, h, j), ‘definite article’ (15g), ‘demonstrative’ (15d), ‘possessive NP’ (15b), ‘wh-word’ (15o), ‘possessive pronoun’ (15i), and ‘quantifier’ (15c).
- QUANTIFYING EXPRESSION: the presence (15l, m, n) or absence of a quantifying expression, as listed in Table 1, before the numeral but within the same phrase.
- N1 NUMBER: whether the numeral itself is ‘one’ (15a, j) or more than one.
- NUMERAL SIZE: values are ‘small’ for the numerals one to nine, irrespective of whether they are written as words or digits, and ‘large’ for all other numerals (15b, c, d, h, i, k, l). This variable was included to check whether any apparent orthographic effect of using digits was an artefact of the stylistic tendency to use words for numerals less than ten, and digits otherwise.
- ADJECTIVE: the presence (15j, m, n) or absence of an adjective or adjective phrase between N1 and N2. The initial corpus searches did not distinguish between the cases exemplified in (15j, m, n), where the adjective modifies N2, and tokens of the MEASUREMENT AS MODIFIER construction, such as *his 15-year old daughter* (BNC ANJ 485), where the adjective forms a constituent with N1 rather than N2. However, the MEASUREMENT AS MODIFIER type were as far as

possible removed from the dataset during manual cleaning, so the remaining adjectives should all be modifiers of N2.

- N2 NUMBER: possible values are ‘singular’, ‘plural’, or ‘uncountable’. This variable was coded in a series of steps. From the CELEX lexical database (Baayen et al. 1995), we extracted the countability information for lemmas classed as nouns. Each lemma is listed in the database with information about whether it occurs as a count noun, and whether it occurs as a non-count noun and it is possible for a lemma to be listed both as count and non-count. In our dataset, if the wordform of N2 differed from the lemma form, and the lemma was listed as potentially countable, we coded N2 as plural. In a second step, we used information about the determiners. If the determiner was *a, an, another, each, every, any one, just one, more than one, my one*, or *such a*, and the lemma was listed as potentially countable, we coded N2 as singular. If there was no determiner, or the determiner was *some*, we coded N2 as uncountable provided the lemma was listed as potentially uncountable, and N2 had not already been marked plural. This left a set of 4,178 lines for which the number of N2 had not been coded, including 514 lines where the N2 lemma was missing from CELEX. These lines were inspected manually and coded for N2 number by either the second author or one of the student assistants.

- N2 COUNTABILITY: whether the final noun is countable or uncountable (15b, e, i, k, l, m, n), based on N2 number as described above.

3.5 Length

To check for possible effects of the number of words in our time measurement expressions or the phrases containing them, we included the following length variables:

- PHRASE LENGTH: the number of words in the full noun phrase containing the time-measurement expression. Hyphenated combinations of numeral plus N1 were counted as two words.
- NP PRE N1 LENGTH: the number of words in the phrase up to but not including the time-noun.
- NP POST N1 LENGTH: the number of words in the phrase following the time-noun.
- N2 LENGTH: the number of nominal constituents of N2, either 'one', where the hit included a simple N2, or the number of nominal compound constituents where N2 is a compound. This is the same as NP post N1 length, minus the number of adverbs and adjectives intervening between the time-noun and final noun.

3.6 Frequency

To enable us to check for any usage-based effects on the distribution of the different forms, we included various frequency variables calculated from the BNC frequency list provided by Sebastian Hoffmann (Hoffmann et al. 2008).

- N2 FREQUENCY: We calculated three different frequencies for each noun: the frequency of the wordform across all parts of speech, the frequency of the wordform tagged as a noun, and the lemma frequency of the noun, i.e. the number of BNC tokens with the singular or plural form tagged as any kind of noun.
- NUM-N1 FREQUENCY: this is the number of tokens in our dataset with a given combination of numeral plus N1, with N1 lemmatised. Two versions of this frequency were calculated: one based on the orthographic representation of the numeral and the other based on the phonological representation. In the first case, *one-hour* would be counted separately from *1-hour*, whereas for the second figure they would be combined.
- NUM-N1-N2 FREQUENCY: this is the number of tokens in our dataset with each unique combination of numeral, N1 lemma and N2

lemma, e.g. ‘three’ + ‘hour’ + ‘work’. Again, two versions were calculated: one based on the orthographic representation and another based on the phonological representation.

- N1-N2 FREQUENCY: the number of tokens in our dataset with each unique combination of N1 lemma plus N2 lemma, e.g. ‘hour’ (N1) + ‘work’ (N2).

All frequencies were logarithmised before being entered into our analyses to guard against disproportionate effects of extreme values.

3.7 Semantics

In order to explore the possible semantic correlates of the different form variants under investigation, we coded the final noun of each hit for its semantic category. Our classification was based on Lyons’ (1977: 442-7) ontological typology, consisting of a threefold distinction between first, second, and third order entities. According to Lyons (1977), ‘first order entities’ are entities that can be located in space, such as a cake or a table; ‘second order entities’ can be located in time and can be said to take place, like a speech or a wedding; ‘third order entities’ are mental expressions that can be evaluated in terms of their truth, such as a belief or a thought. Additionally, we used Hengeveld and Mackenzie’s (2008: 136) expanded

typology of semantic categories, as they make further discriminations that were required to account for all the distinctions found in our dataset. In this typology, 'individual', 'state-of-affairs', and 'propositional content' correspond to Lyons' first, second, and third order entities, respectively. To these categories, Hengeveld and Mackenzie (2008) have added 'property', 'location', 'time', 'episode', 'manner', 'reason', and 'quantity'. Examples of nouns illustrating these different semantic categories are shown in Table 2.

Table 2: Semantic coding of N2, based on Lyons (2007); Hengeveld and Mackenzie (2008)

Category	**Example**
first order entity (individual)	*chair*
second order entity (state-of-affairs)	*meeting*
third order entity (propositional-content)	*idea*
property	*colour*
quantity	*litre*
location	*top*
time	*week*
episode	*incident*
manner	*way*

reason	*reason*

We extracted the final noun from each token in our dataset, lemmatised them, deleted duplicates, and then classified the resulting set of types in terms of the semantic categories described above. The classification of each noun out of context was not always straightforward as, more often than not, the same noun can refer to different semantic categories in different contexts. Bearing this in mind, we used two different criteria. Firstly, as certain affixes (e.g. *-ation, -age*, *-ing*) are generally associated with the expression of actions, processes, events, or activities (Hengeveld & Mackenzie 2008; Bauer, et al. 2013 for further affixes), derived nouns with any of these suffixes, e.g. *acclimatization*, were classed as expressing second order entities. Secondly, we checked dictionary definitions, specifically the definitions provided by the *Cambridge dictionary online*. For example, if a noun was defined as 'the act (action, event, process) of …', or 'the state of …', then this noun was labelled as a second order entity (state-of-affairs), as actions, events, processes, and states are all things that can be located in time. Thus, *arrest* was analysed as a second order entity because it was defined as 'the act of arresting someone'. This was extended to nouns that were defined in terms of other nouns that were themselves defined as an act or a state. For example, *abuse* is defined as 'the use of something in a way that is harmful or morally wrong', where *use* is in turn

defined as ‘the act of using something, or a period of time when something is being used or can be used’.

Because most words are polysemous, different definitions of the same word, or even a single definition, can correspond to more than one type of entity. For example, *absence* is defined as ‘the state of not being somewhere, or a period in which you are not somewhere’. The first part of this definition denotes a state, i.e. an entity located in time, whereas the second part denotes time itself. In such cases, the relevant hits were checked in context, either by the second author or one of the student assistants, and coded at token level for entity type.

Of the ten categories listed in Table 2, we used all except ‘episode’ and ‘reason’. For ease of analysis, we created eight binary variables, as follows:

- ORDER1
- ORDER2
- ORDER3
- PROPERTY
- QUANTITY
- LOCATION
- TIME
- MANNER

Each of these variables has the values ‘yes’ or ‘no’, corresponding to whether or not the noun has a reading that falls into the relevant category.

4. Model 1: Apostrophe use with S-form N1

4.1 Methodology

In order to investigate the factors that influence the use of the apostrophe in time-measurement expressions, we selected only that part of our data that had originated from the written portions of the BNC. This was to avoid the risk that the analysis would be disproportionately affected by the orthographic preferences of the transcribers of the spoken corpus. The data from the written BNC consisted of 16,555 tokens. We then further reduced this set to include only those 3,497 tokens in which the time word was spelt with a final *s*, with or without an apostrophe. To be able to include author type as a variable in the model, we subsequently removed an additional 313 tokens for which this information was not provided by the corpus.

We carried out logistic generalised mixed effects regression analysis using the lme4 package in R (Bates et al. 2015). Logistic regression involves a binary dependent variable, in this case the presence or absence of an apostrophe. To check whether particular lexical items are associated with

use or omission of the apostrophe, as suggested by Biber et al. (1999: 293), we included a random intercept for N2 lemma; in other words, we allowed the baseline probability of apostrophe use to vary according to the lemma of N2. However, random effects can only meaningfully be included in models when there are repeated measures in the data. For this reason, we excluded examples with a unique final noun, further reducing the number of tokens to 2,951 (for an introduction to mixed-effects models, see Baayen et al. 2008). To select the best models for our data we used the Akaike Information Criterion (AIC, Akaike, 1974), a measure that evaluates alternative models of a dataset on the basis of both their goodness-of-fit and their complexity.

Several of the numerical predictors in our data were highly correlated with each other. For example, all three versions of N2 frequency were highly correlated, as were the orthographic and phonological versions of the various string frequencies. Since including highly correlated predictors in statistical models can produce misleading results, we took steps to reduce the level of collinearity. For each set of correlated measures, we created models with just one of the variables, and then selected the variable that produced the lowest value of AIC in its individual model to be included in our subsequent modelling. As a result of this procedure, for the model of apostrophe use, we included the wordform frequency of N2, the orthographic versions of num-N1 frequency and num-N1-N2 frequency, the number of nouns in N2, and the length of the noun phrase up to N1.

Exclusion of the other numerical variables reduced collinearity to an acceptable level, as indicated by a condition number of 20.75[3] (languageR package, Baayen and Shafaei-Bajestan 2019).

Further inspection of the data for sex of author revealed that, for the majority of items in our dataset, the sex of the author was not recorded in the BNC, and we therefore excluded this variable from consideration. We also excluded the semantic variable 'manner', since none of the tokens in this sub-set of the data had been coded with that entity type. In our initial modelling we included all the other meta-textual and semantic variables outlined in Section 3 as fixed-effect predictors, as well as the orthographic variables 'hyphen' and 'numeral form'. We likewise included all the remaining morphosyntactic variables except for 'N1 form', which was irrelevant to this dataset from which the bare-form tokens had been excluded. The full list of predictors included in our modelling is shown in Table 3.

Table 3: Predictors included in the modelling of apostrophe use

Predictor variables	
Metatextual categories	
Derived text type	Type of author
Orthography	

[3] According to Baayen (2008), condition numbers of 30 or more may indicate potentially harmful collinearity.

Hyphen	Numeral form
Morphosyntax	
Preposition	Preposition type
Premodifier	Determiner
Determiner type	Quantifying expression
N1 number	Numeral size
Adjective	N2 number
N2 countability	N2 lemma (random intercept)
Length	
NP pre N1 length	N2 length
Frequency	
Log N2 frequency (word-form)	Log num-N1 frequency (orth.)
Log num-N1-N2 frequency (orth.)	
Semantics	
Order1	Order2
Order3	Property
Quantity	Location
Time	

For model optimisation we used a forward selection procedure based on AIC. We started by creating an individual model for each predictor plus the random effect for N2 lemma. We then selected that variable that produced the lowest value of AIC in its individual model for inclusion in our subsequent models. We proceeded in this way, adding further variables until

adding the next remaining variable with lowest AIC in its individual model did not reduce the AIC by at least 2. Having selected the predictors for our model in this way, we then checked whether any interactions between them further reduced AIC, and finally checked that the random effect still improved the model.

4.2 Results and discussion

The final model, after model optimisation, is shown in Table 4 and represented graphically in Figures 1 and 2.

Table 4: Final mixed effects model for apostrophe use

Random effects:					
Groups	Name	Var.	Std.Dev.		
N2Lemma	Intercept	0.6721	0.8198		
Number of obs: 2,951, groups: N2Lemma, 189					
Fixed effects:					
	Estimate	Std. Err.	z-value	Pr(>|z|)	
(Intercept)	0.99176	0.34692	2.859	0.00425	**
TextType WA	-0.45586	0.24445	-1.865	0.06220	.
TextType WP	-0.54668	0.22394	-2.441	0.01464	*
TextType WO	-0.63589	0.23805	-2.671	0.00756	**
TextType WN	-1.18090	0.25306	-4.666	3.06e-06	***
TextType WU	-1.96364	0.28617	-6.862	6.80e-12	***
Author Multiple	-0.28379	0.13002	-2.183	0.02906	*
Author Corporate	-0.83728	0.17367	-4.821	1.43e-06	***
NumeralForm Word	0.30722	0.10714	2.867	0.00414	**

NPPreN1Length	0.10631	0.05261	2.021	0.04331	*
N2Length	-0.60635	0.15678	-3.868	0.00011	***
N2Countability	-0.13639	0.28384	-0.481	0.63086	
Preposition	0.78376	0.22057	3.553	0.00038	***
LogNumN1Freq	0.16888	0.03875	4.358	1.31e-05	***
N2Countability:Prep	-1.30243	0.43439	-2.998	0.00271	**
Prep:logNumN1Freq	-0.09162	0.04560	-2.009	0.04452	*

Signif. codes: 0 '***', 0.001 '**', 0.01 '*', 0.05 '.'

The top section of the table shows the random effect for N2 lemma, which remains significant in the final model. The lower half of the table shows the significant fixed effects. There are significant main effects for derived text type, author type, the number of constituent nouns in N2, the number of words before N1, and the orthographic form of the numeral. There are also significant interactions of preposition use with N2 countability and the frequency of num-N1. A positive coefficient in the first column indicates that the relevant value of the predictor is associated with increased probability of apostrophe use, whereas a negative coefficient indicates that the value is associated with lower probability of apostrophe use.

In Figures 1 and 2, the labels on the vertical axis of each plot indicate the probability that an apostrophe will be used, and the horizontal axes represent the values of the significant predictors. For the categorical variables, the dots on the graphs indicate the mean estimated likelihood of an apostrophe being used in contexts that have the pertinent value of that variable. For the numeric variables, the graphs show regression lines. The

whiskers around the dots and the grey shading around the regression lines represent the 95% confidence intervals. In order to show the effect of each predictor in turn, the other predictors are adjusted to their reference level, for categorical variables, or to the mean for the numerical predictors. The reference level for each categorical variable is the left-most value in the relevant plot. In other words, the model shows the effect of independently varying each predictor in a situation where each of the other predictors has the value shown to the left of its x-axis. The reference level for the random effect is its most frequent value, which is *time.*

Figure 1 shows the significant main effects. The strongest effects are those of text and author type. The top left-hand plot shows that an apostrophe is less likely to be used in newspapers than in other kinds of publication, and least likely of all in unpublished writing. The second plot on the top line shows the effect of author type. Apostrophes are less likely to be used in corporate publications than in those with a named author or authors. The plots on the second row of Figure 1 show significant effects of the lengths of different parts of the expression. The left-hand plot shows that an apostrophe is less likely to be used when N2 is a compound than when it is not a compound, and that an apostrophe becomes increasingly unlikely the more compound constituents there are. The right-hand plot on the second row shows that the more words occur in the noun phrase before N1, the more likely it is that an apostrophe will be used. These two length

effects indicate that writers are more likely to use an apostrophe when the opportunity to do so arises later in the noun phrase, suggesting that the punctuation might sometimes have a sort of pause function. The final plot, on the bottom row of Figure 1, shows that, all other things being equal, an apostrophe is more likely to be used when the numeral is written in words than when it is written using digits.

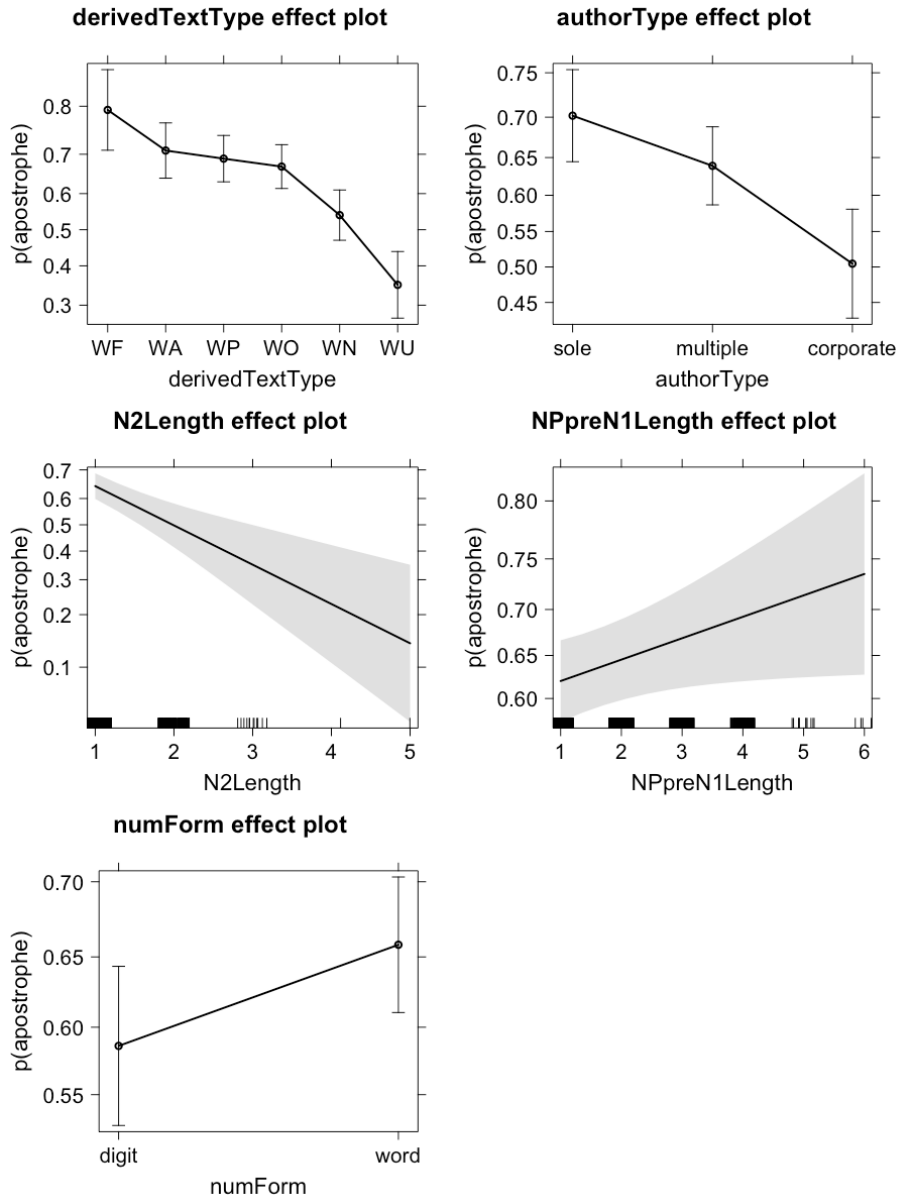

Figure 1: Partial fixed main effects in the final model for apostrophe use

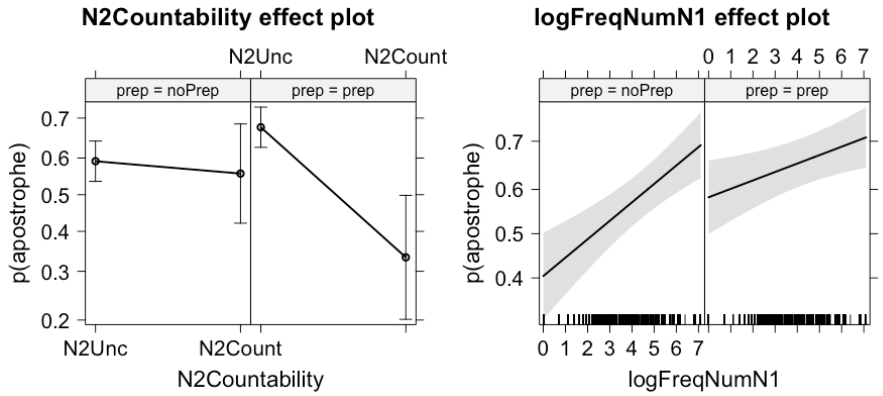

Figure 2: Significant interaction effects in the model for apostrophe use

Figure 2 shows the two interaction effects in the final model for apostrophe use, both of which involve the presence or absence of a preposition. On the left-hand side we see the interaction of N2 countability with preposition use and on the right-hand side we see the interaction of the frequency of num-N1 with preposition use. In other words, both N2 countability and num-N1 frequency relate differently to apostrophe use when the time-measurement expression occurs in a preposition phrase, than when it occurs only in a noun phrase. The right-hand plot in Figure 2 shows that writers are more likely to use an apostrophe with more frequent num-N1 combinations, and that this effect is most marked when the expression is not part of a preposition phrase. In the left-hand plot of Figure 2, it can be seen that N2 countability only correlates with apostrophe use in preposition phrases, and an apostrophe is particularly unlikely to be used when a time-measurement expression occurs in a preposition phrase and N2 is countable.

Inspection of the dataset revealed that only 36 tokens actually have this constellation of properties, of which 29 tokens (83%) belong to the PHRASAL COMPOUND construction. Some examples are shown in (16):

(16) a. *The children were assessed on two occasions, each separated by a **three months interval**.* (BNC ALM 1143)

b. *Thus can a man, in three **five-minutes sessions**, attain a multilateral tan.* (BNC A2J 200)

c. *The conception of a **three-years honours course** incorporating as many as fifteen weeks teaching practice …* (BNC HTK 743)

It is unsurprising that the apostrophe is relatively unlikely to be used in the phrasal compound types because, as discussed in Section 2, these expressions have a different syntactic distribution from the *s*-genitive and are therefore more likely to be perceived as unambiguously common case plural; however, it is unclear to us why this effect should only emerge in prepositional phrases. What is clear is that, for the overwhelming majority of types (99%), there is no correlation between apostrophe use and any semantic or morphosyntactic variable. By far the strongest effects in this model are the effects of text and author types. This suggests that the choice about whether or not to use an apostrophe is largely stylistic and that the

versions with and without an apostrophe may be orthographic variants of the same construction; the effect of numeral form further supports this conclusion.

Turning now to possible lexical effects, Figure 3 shows the variation between N2 lemmas in terms of the baseline probability of apostrophe use. The dots represent the intercept for each noun. The further to the right a dot appears, the greater the tendency of time-measurement expressions with that noun to include an apostrophe. The horizontal lines represent the 95% confidence intervals for these intercepts; if the confidence interval crosses zero, we cannot be entirely confident that the noun in question exhibits any significant bias. It can be seen that, although inclusion of the random effect improves the predictive power of the model overall, in fact quite few of the nouns exhibit a clear tendency in either direction, suggesting that lexical preferences do not play a major role in apostrophe use. The nouns that show a preference for apostrophe use are *notice, experience, pay, gestation*, and *jail*, while *incubation, difference, walk, difficulty, interval*, and *time* have the opposite tendency.

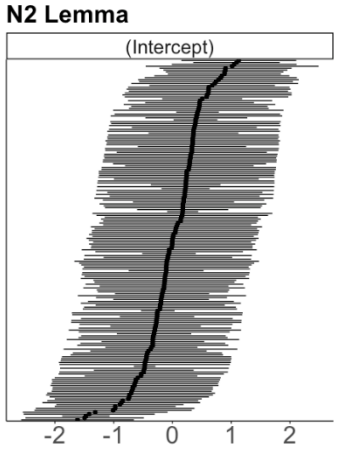

Figure 3: By-N2 random intercepts in the final model for apostrophe use

On the face of it, these findings differ from those of Biber et al (1999: 293), who report that the apostrophe tends to be used regularly with *time* but omitted with *imprisonment*. However, it should be remembered that the random effect for N2 lemma in our model shows tendencies for lexical items after controlling for all other effects. To check whether this could explain the difference between our result and that of Biber et al. (1999: 293), we therefore also checked how often *time* and *imprisonment* occur with apostrophes in our raw data. It turns out that about 58% of tokens with S-form 'time' have an apostrophe, while about 80% of tokens with S-form 'imprisonment' have an apostrophe. In other words, even looking at the raw data, our results go in the opposite direction to those of Biber et al. (1999: 293): a reminder, if one were needed, that results based on a single corpus should be interpreted with caution.

The cautionary tale of the previous paragraph notwithstanding, it is striking that of the eleven morphosyntactic and seven semantic variables entered into this analysis as fixed effects, only two – preposition use and countability of N2 – emerge as significant, and only in interaction with one another. Figure 4 shows the distribution of these variables across the whole dataset of S-form N1 types in the written corpus. The figure is type-based in the sense that tokens were excluded if the whole phrase containing the time-measurement expression duplicated a phrase already in the data. Since we do not need to exclude tokens with a unique N1 for this plot, nor those without author information, we are left with 2,599 types.

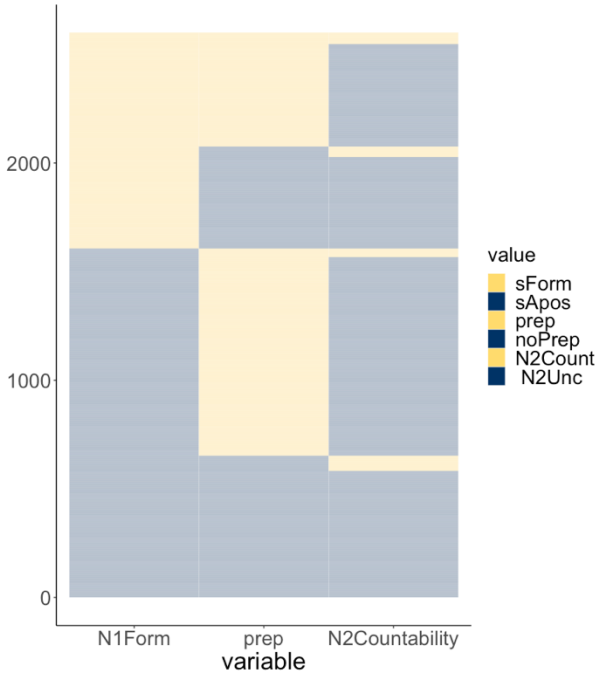

Figure 4: Co-occurrence of significant morphosyntactic features in the data with S-form N1

Each column in the figure represents one of the variables, namely apostrophe use, presence or absence of a preposition, and countability of N2. In each column, the dark shading represents types that have one value of that variable and the light shading represents types that have the other value. Each type is represented by a horizontal line, so that the figure overall represents the extent to which the various features co-occur. The top half of the graph, with light shading in the first column, represents types without an apostrophe and the bottom half of the graph, with dark shading in the first column, represents types with an apostrophe. In both areas, the distribution of the other two variables is similar. In the second column, we see that S-forms both with and without apostrophe occur both with and without prepositions, although a preposition is relatively more likely with an apostrophe. In the third column we see that, for all four possible combinations of N1-form and preposition use, a small minority of N2 types are countable. Overall, no set of features represented by the values of these variables always cluster together. We take this to indicate that, apart from a very few phrasal compounds, the forms with and without an apostrophe do not belong to different constructions but rather they are more helpfully viewed as orthographic variants of the same construction, with the presence or absence of an apostrophe being determined by stylistic and contextual factors.

5. Model 2: Use of the S-form

5.1 Methodology

Having concluded that the presence or absence of an apostrophe in time-measurement expressions is not a criterion for distinguishing different constructions, we now turn our attention to the question of what influences the choice between S-form N1 with or without apostrophe (*10 days free trial, six months' interest*) as opposed to a bare-form N1 (*11 year reign*). Since this concerns a phonological distinction, we started with the full set of 17,591 tokens, including those from both the spoken and written sections of the corpus and all forms of N1. After removing those without author information, we were left with 15,266 tokens.

We again carried out logistic generalised mixed effects regression analysis using the lme4 package in R (Bates et al. 2015), using the same optimisation strategy as described for Model 1. The dependent variable in our model was the form of N1: S-form or bare-form. To check whether particular lexical items are associated with use of the S-form, we included a random intercept for N2 lemma, allowing the baseline probability to vary according to the lemma of the final noun. For this reason, we excluded

examples with a unique final noun, reducing the number of tokens to 14,620.

We took steps to reduce the level of collinearity amongst our numerical predictors by selecting a subset of these predictors to include in our modelling, also using the same procedure as described for Model 1. The variables included in our models were the same as those shown in Table 3, except that the frequency of N2 wordform tagged as a noun replaced the frequency of the wordform across all parts of speech, and the frequency of N1N2 replaced frequency of num-N1N2. We also added the variable 'manner', since this value of entity type was used for a few tokens in the full dataset.

5.2 Results and discussion

The final model after optimisation is shown in Table 4 and graphically in Figure 5 (stylistic and orthographic effects) and Figure 6 (morphosyntactic and semantic effects). The y-axis of the plots is labelled with the probability of N1 having an S-form. This probability is always low, because the bare-form is much more frequent, accounting for 76% of tokens in the dataset.

Table 5: Final mixed effects model for S-form use

Random effects:			
Groups	Name	Variance	Std.Dev.

N2Lemma	Intercept	3.537	1.881	

Number of obs: 14,620, groups: N2Lemma, 790

Fixed effects:

	Estimate	Std. Err.	z-value	Pr(>|z|)	
(Intercept)	0.50870	0.66260	0.768	0.442647	
N2Number Singular	-5.53950	0.29850	-18.558	< 2e-16	***
N2Number Plural	-5.91242	0.31534	-18.749	< 2e-16	***
N1Number Plural	4.13617	0.51851	7.977	1.50e-15	***
Quant.Expression	2.48700	0.34540	7.200	6.00e-13	***
TextType WN	-0.22316	0.44388	-0.503	0.615151	
TextType WP	-0.60144	0.38941	-1.544	0.122471	
TextType WO	-0.96288	0.41338	-2.329	0.019845	*
TextType WU	-0.97004	0.46675	-2.078	0.037682	*
TextType WA	-1.17370	0.43624	-2.691	0.007134	**
TextType SC	-1.45383	0.51688	-2.813	0.004912	**
TextType SO	-2.14520	0.41903	-5.120	3.06e-07	***
Author Multiple	-0.94796	0.21348	-4.440	8.98e-06	***
Author Corporate	-1.40291	0.28123	-4.989	6.08e-07	***
Order2	-0.80000	0.24177	-3.309	0.000937	***
PrepType Other	0.04732	0.16521	0.286	0.774563	
PrepType Time	0.59415	0.19595	3.032	0.002429	**
Premodifier	-2.99435	0.85409	-3.506	0.000455	***
Determiner	-1.23906	0.23838	-5.198	2.02e-07	***
NumeralForm Word	0.45246	0.17083	2.649	0.008081	**
Hyphen	-5.86429	0.53071	-11.050	< 2e-16	***
PreMod:Det	4.27671	0.88756	4.819	1.45e-06	***
NumForm:Hyphen	1.39732	0.59437	2.351	0.018725	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

As with the model for apostrophe use, there are significant effects of text and author type, shown in the plots on the top row of Figure 5. The S-form

is most likely to be used in fiction and verse and least likely in spoken language. The S-form is also most likely to be used by sole authors and least likely to be used by corporate authors. It appears that author types that favour apostrophe use also favour the S-form. In addition, the form of N1 interacts with other aspects of orthography. The interaction plot on the bottom of Figure 5 shows that hyphens are much less likely to be used when N1 has the S-form, especially if the numeral is written as a digit.

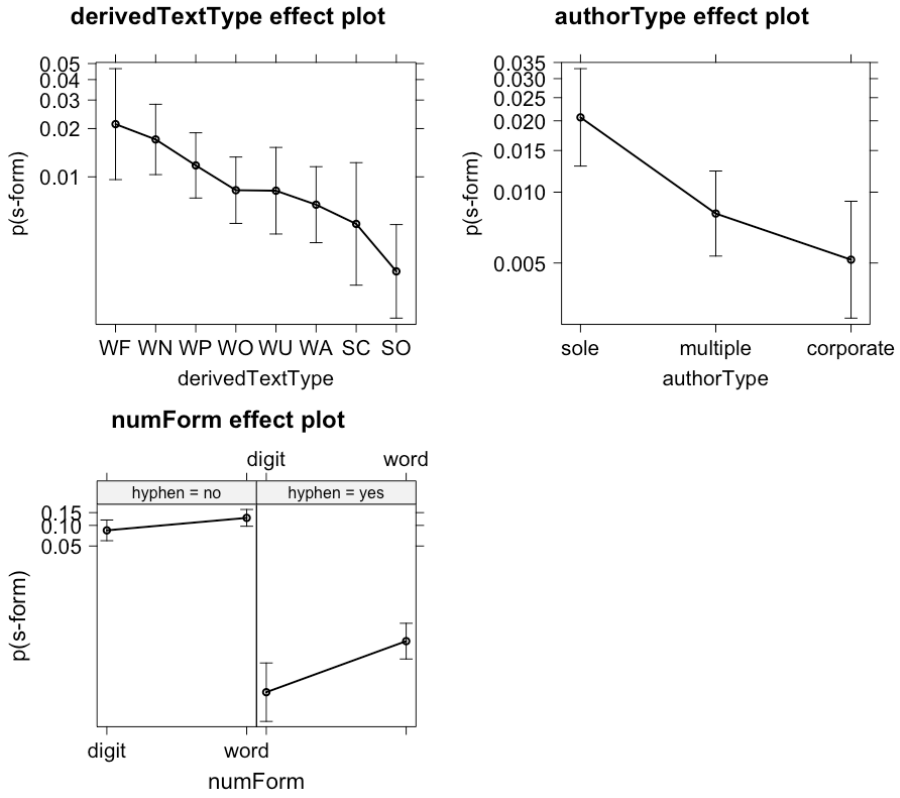

Figure 5: Partial effects of stylistic and orthographic variables in the final model for S-form use

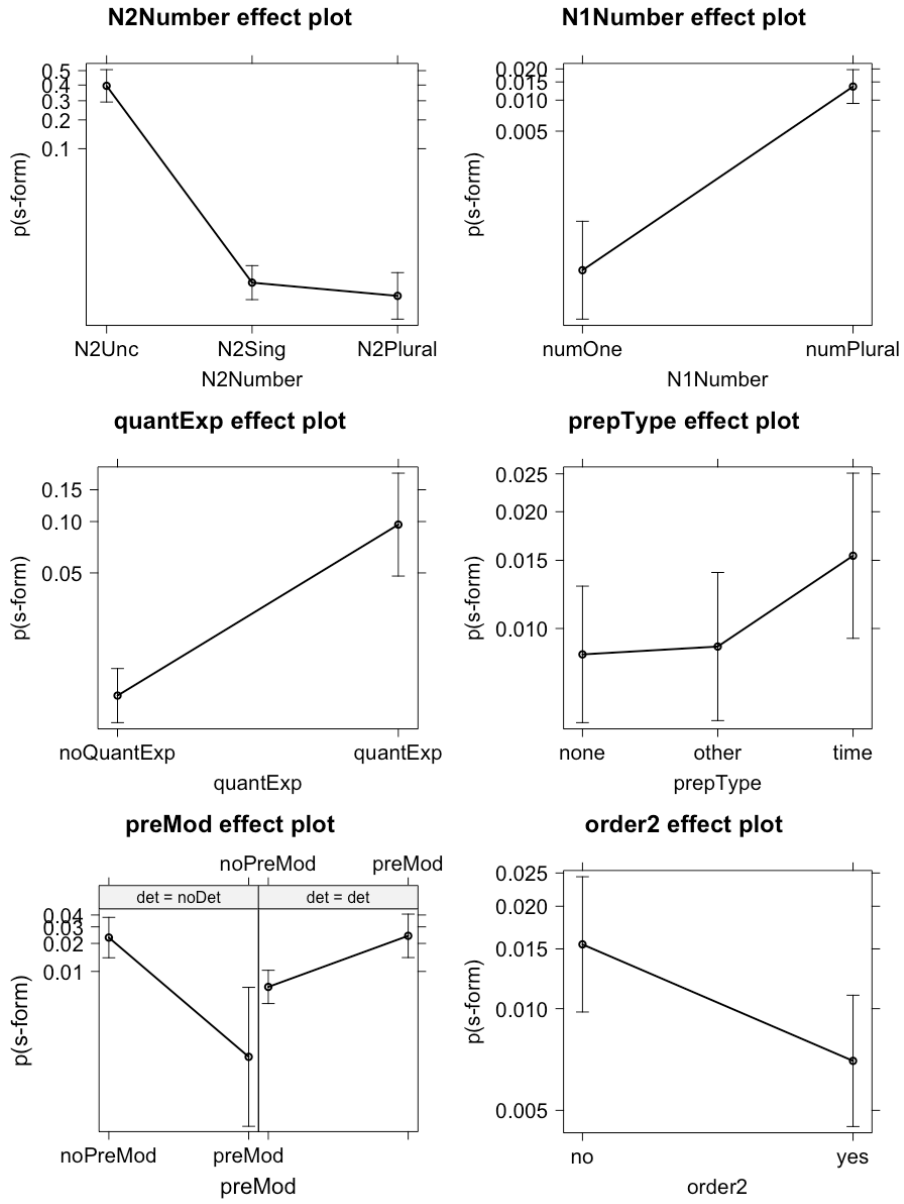

Figure 6: Partial effects of morphosyntactic and semantic variables in the final model for S-form use

In contrast to the model for apostrophe use, in this model we see much stronger and more extensive effects of the structure of the expressions. Figure 6 shows the effects of the semantic and morphosyntactic variables in the final model. It can be seen that by far the strongest effect is the number of N2. The S-form is far more likely to be used when N2 is uncountable than when it is countable, either singular or plural. There is also an effect of N1 number: the S-form is less likely to be used when the numeral is 'one' than when it is greater than one.

The two plots on the middle row and the interaction plot on the bottom row of Figure 6 show the correlations between the form of N1 and the presence or absence of other elements in the phrase. The S-form is more likely to be used when the phrase includes a quantifying expression and when it includes a time preposition. Remember that both the quantifying expressions and prepositions of time may enter into constituents with the num-N1 element of the phrase, so the correlation of their presence with S-form use suggests that the num-N1 element tends to be more complex in expressions with the S-form. This is consistent with the idea that time-measurement expressions with the S-form may be noun phrases whereas those with the bare-form might be simpler nominals.

The left-hand plot on the bottom row of Figure 6 shows the interaction effect of pre-modifiers and determiners on S-form use. The left-hand panel shows the effect of pre-modifier presence when there is no

determiner, and the right-hand panel shows the effect of pre-modifier presence when there is a determiner. It can be seen that the probability of S-form use is greatest when both pre-modifier and determiner are absent or when they are both present. Like the interaction between preposition use and N2 countability in the model of apostrophe use, this effect is driven by a very small subset of the data since the great majority of tokens with S-form N1 do not have a determiner. There are only 289 tokens with both S-form N1 and a determiner, representing just 12% of S-form tokens and fewer than 3% of tokens with a determiner. However, of these 289 tokens, some 56 (19%) also have a premodifier, compared with only 11% of tokens with bare-form N1 and determiner. A glance at the data shows that when a premodifier occurs with a determiner and S-form N1, the determiner and premodifier are likely to belong to N1 rather than N2. For example, we interpret (17a) as [[*the first two days*] *golf*] and (17b) as [[*an extra 5 years*]*optional cover*].

(17) a. ***the first two days' golf*** *featured a Pro/Am event …*
(BNC HPC 465)

b. *5 years, with* ***an extra 5 years optional cover****.*
(BNC ASD 1063)

(18) a. *… his rent would more than double under a new **20-year lease*** (BNC K55 4423)

b. *The Bill also imposes an automatic **one-year driving ban** …* (BNC HHX 7081)

In contrast, in the relatively few cases where a premodifier occurs with a determiner and bare-form N1, the determiner and premodifier are likely to belong to N2, as exemplified in (18a), [*a* [*new* [*20-year lease*]]] and (18b), [*an* [*automatic* [*one-year driving ban*]]]. This is again in keeping with the hypothesis that time-measurement expressions with S-form N1 are noun phrases whereas those with the bare-form are nominals.

Finally, there is a small effect of the entity type of N2 on the probability of S-form use, with the bare-form being slightly more likely when N2 is a second order entity. In fact, second order entities are by far the most frequent semantic class across all types in our data. This is not surprising since second order entities are, by definition, those which have a duration and all the expressions in our dataset were selected because num-N1 represented a duration or related quantity.

Regarding possible word-level effects, as with Model 1, a random effect for N2 lemma improves overall model performance, but very few nouns show a significant deviation from the mean once all the other effects are taken into account. In other words, the effects of more general

properties, especially the countability of N2, are stronger than the more specific effects of individual lemmas. Nevertheless, a few nouns do show significant attraction for one form or the other. The lemmas that show strongest association with bare-form N1 are *care*, *healing*, *survival*, *monitoring*, *retention*, *mortality*, and *acidity*. These are all uncountable nouns that nevertheless occur mostly with bare-form N1, usually in the context of medical or scientific texts and in the case of *care* almost exclusively collocated with '24-hour'. Examples are shown in (19):

(19) a. *But he suffers from motor neurone disease and needs* ***twenty-four hour care****.* (BNC K1C 1358)

b. *Comorbidity emerged as the best predictor of* ***2-year survival****.* (BNC HWU 1059)

The lemmas with strongest attraction to the S-form include e.g. *worth*, which occurs exclusively with the S-form, and *wonder*, which occurs with the S-form in the context of *nine-days' wonder*, despite being coded as countable in that context.

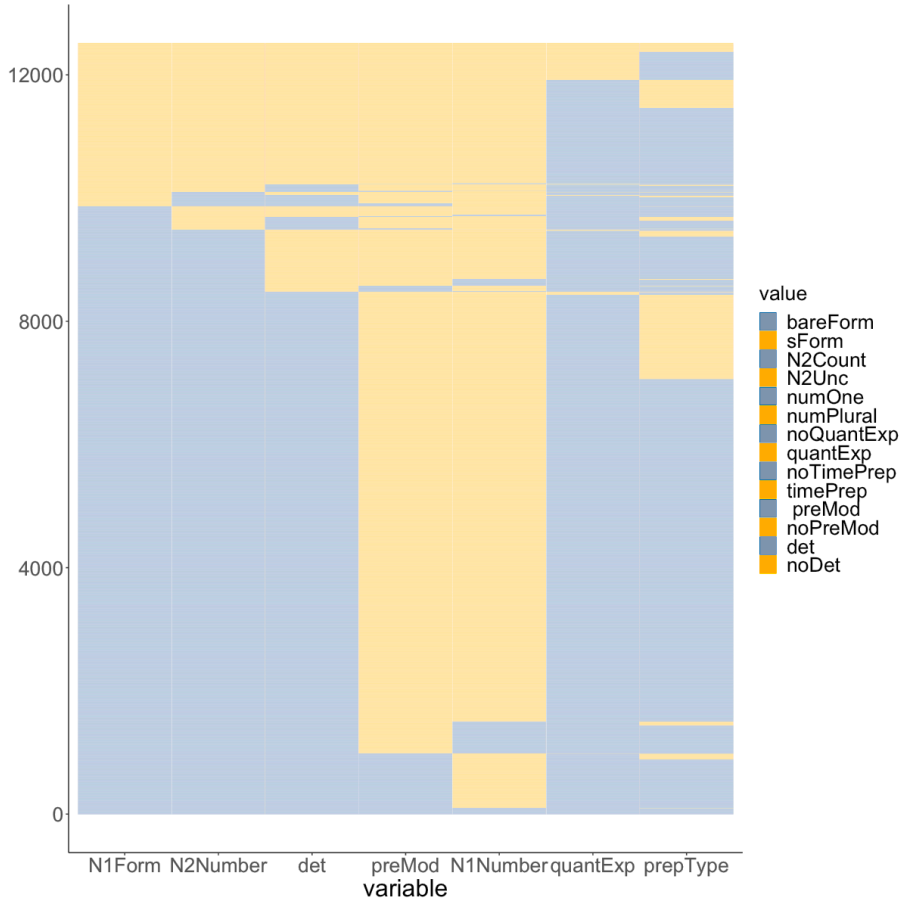

Figure 7: Co-occurrence of morphosyntactic features associated with use of S-form N1

To what extent can the structural tendencies identified in the regression model be taken to indicate that it would be useful to recognise different constructions? To explore this issue, we again plotted the co-occurrence of the various morphosyntactic features, as shown in Figure 7. This plot, based on 12,516 phrase types after the exclusion of duplicates, shows the extent to

which the statistically significant morphosyntactic features identified above – as opposed to stylistic or orthographic features – tend to cluster together.

In Figure 4, we took the similar patterns in the top and bottom halves of the graph to indicate that the presence or absence of an apostrophe did not merit separate constructional status. In contrast, in Figure 7, the large light rectangle in the top left corner shows that in types with S-form N1, with or without an apostrophe, there is a clustering of features distinct from types with bare-form N1. We conclude that these clusters of properties represent two different constructions. There is a particularly strong tendency of S-form N1 to occur not only with plural numerals but also with uncountable N2. These expressions are unlikely to include a determiner of any type, or a premodifier, but are relatively likely to include a 'quantifying expression', compared to similar expressions with bare-form N1. As far as we have been able to ascertain, the strong association of S-form N1 with uncountable N2 has not been discussed in previous literature and appears to have gone unnoticed. Although Biber et al. (1999: 293) mention that with countable N2 the time-noun is normally in the singular, they do not follow through with the corollary that S-form time-nouns are associated with uncountable N2. This association does not follow, for example, from general formal properties of genitive constructions, which occur freely with countable heads. We call this construction, with S-form N1 and normally uncountable N2, the TIME-MEASUREMENT construction. In contrast, the

TIME-MEASUREMENT COMPOUND construction is characterised by bare-form N1 and normally countable N2. In the next section, we will explore these two constructions in more detail, using examples from our dataset.

6. Characteristics of the two constructions

6.1 Overview

We set out to investigate whether the different orthographic forms of the time word (N1) in time-measurement expressions are indicative of different constructions. We characterised these forms as the bare-form (e.g. *week*), the S-form with apostrophe (e.g. *weeks'* or *week's*), and the S-form without apostrophe (e.g. *weeks*). In Section 4, we showed that apostrophe use is not consistently correlated with any other aspect of form or meaning that would justify postulating two constructions with S-form N1, one with apostrophe and the other without apostrophe. On the other hand, in Section 5, we showed that expressions with the bare-form tend to differ in several important respects from expressions with the S-form, with or without apostrophe. We concluded that these two forms of N1 are characteristic of two different constructions: the TIME-MEASUREMENT COMPOUND construction and the TIME-MEASUREMENT construction, respectively.

Table 6: Typical properties of the Time-measurement construction compared with the Time-measurement compound construction

Time-measurement compound construction	**Time-measurement construction**
Bare-form N1	S-form N1, with or without apostrophe
Countable N2	Uncountable N2
Preceded by a determiner	No determiner
Num-N1 hyphenated	No hyphen

On the basis of the analysis presented in Section 5, we can identify a cluster of properties that distinguish typical exemplars of the TIME-MEASUREMENT COMPOUND construction from typical exemplars of the TIME-MEASUREMENT construction. These are summarised in Table 6. In our dataset, 96% of tokens with bare-form N1 have a countable N2 and 87% also have a preceding determiner; this combination of properties characterises the central representatives of the TIME-MEASUREMENT COMPOUND construction. Conversely, 94% of tokens with S-form N1 have an uncountable N2, and 91% also lack a determiner, properties that characterise the central members of the TIME-MEASUREMENT construction. Orthographically, the two constructions are distinguished not only by the presence or absence of *s*, but also by hyphenation: 73% of tokens with bare-form N1 and countable N2 have a hyphen between the numeral and N1, whereas fewer than 1% of

tokens with S-form N1 and uncountable N2 have a hyphen. However, despite these strong tendencies, it will be clear from the percentages, as well as from the patterns in Figure 7, that the clustering of properties is not categorical. Rather, between the typical time-measurement construct and the typical time-measurement compound construct, it is possible to find every combination of relevant properties, and the most frequent combinations therefore represent only areas of greater density in a sea of variation. In the following paragraphs we will start by looking at some examples of the most frequent types before considering some of the variants and finally arriving, in Section 7, at definitions of the two constructions.

6.2 Central exemplars

Typical examples of the TIME-MEASUREMENT construction and the TIME-MEASUREMENT COMPOUND construction, i.e. with the morphosyntactic properties summarised in Table 6, are shown in (20) and (21) respectively.

(20) a. *In **ten years' time** you'll know why you were right to send off this coupon.* (BNC CFT 97)

b. *The course is for **four years' duration**.* (BNC B3C 1912)

c. *Alumasc Systems are superbly engineered by a British company which has* ***over 40 years experience****.*
(BNC CFT 2554)

d. *Of the five other FIS officials accused, one was sentenced to* ***six years' imprisonment*** *and four to four years each.*
(BNC HLM 2317)

(21) a. *Most record company agreements are on the basis of a* ***one-year period*** *plus a number of options.* (BNC A6A 2143)

b. *The May 1968 events in France triggered a* ***three-week general strike****.* (BNC K8U 1327)

c. *The agreement … officially ended the* ***11-year civil war*** *which had cost 75,000 lives.* (BNC HLF 856)

d. *Paul Loughlin has agreed a new* ***two year contract****.*
(BNC K3P 240)

The most important difference between the two constructions lies in the semantics of N2, which is linked to countability and constrains the meaning of the whole expression. Although the relationship between grammatical countability and meaning is complex, countable nouns usually represent individual entities with clear boundaries, or sets of such bounded entities, while uncountable nouns usually represent entities that are unbounded, at

least in the given context (cf. Jackendoff 1996); there is empirical evidence that the ontological and cognitive distinction between uncountable 'stuff' and countable 'things' is largely reflected in the linguistic distinction between mass and count nouns, even though the mapping is not one-to-one nor consistent across all languages (Lin et al. 2018). In English, noun phrases with uncountable head nouns and no determiner virtually always represent unbounded entities (Jackendoff 1991, 1996). As these are the properties of N2 in typical exemplars of the TIME-MEASUREMENT construction, it follows that N2 represents an unbounded entity in such expressions. And the num-N1 element cannot represent the total duration of N2 since, by definition, an unbounded entity does not have a finite duration. Rather, num-N1 represents a quantity of the unbounded entity represented by N2; in other words, the meaning is pseudo-partitive.

In contrast to the TIME-MEASUREMENT construction, N2 in the TIME-MEASUREMENT COMPOUND construction represents a bounded entity or entities, and this semantic property is reflected in the predominance of countable types. Because N2 is bounded, it is possible for num-N1 in these constructs to represent the duration of N2, as it does in (21). However, this is not the only possible meaning. Although we excluded them from our database (where we were focussing on duration), the corpus also includes expressions that have both the form of a typical time-measurement compound and the semantic property of a bounded N2, but num-N1 does

not represent the duration of N2. In these cases, num-N1 represents an amount of time with some other relation to N2. For example, in (22a), *nine-month* is not the duration of the result, but the period over which the result is calculated, and in (22b), *one-day* is not the duration of the series, but the duration of each cricket match in the series[4].

(22) a. *Soaps to detergents giant Unilever lost 28p to 1081p after disappointing third-quarter figures which pushed the* ***nine-month result*** *from £1.3bn to £1.4bn.* (BNC CEL 307)

b. *West Indies and Pakistan will visit South Africa in Feb 1993 for a* ***one-day triangular series****.* (BNC CU0 134)

In determining the meaning of the TIME-MEASUREMENT COMPOUND construction, we cannot therefore be more specific than to say that the construction has the meaning 'a type of N2 with some relation to time-period num-N1', in line with the meaning of endocentric compounds generally (Booij 2009: 201). The form of the construction itself does not specify the meaning further; rather, the specific relation between N2 and N1 is determined partly by the lexical semantics of the specific N2 and partly by the context in which the construction occurs. Nevertheless, by far the

[4] A one-day triangular series is a series of cricket matches each lasting one day, played between all pair-wise combinations of three teams, usually at international level.

most common relation in the case of time-measurement compounds is that num-N1 is the duration of N2; from our total set of corpus hits, fewer than 6% of hyphenated tokens had a different relation. Furthermore, cases where num-N1 is not the duration of N2 often belong to rather specialised vocabularies, as exemplified in (22a) and (22b).

We turn now to the form and function of N1 in the two constructions. In the TIME-MEASUREMENT construction, we have shown that the presence or absence of an apostrophe on the S-form N1 does not generally signal a difference in meaning, but we have not yet addressed the nature of this S-form with optional apostrophe. The tendency of the construction to occur with cardinal numerals greater than one might suggest that the S-form represents a common case plural. However, examples like (23a-c) constitute evidence against a common case analysis, since when the numeral is 'one', the common case would not have the S-form.

(23) a. *All female staff aged 20 years or more, subject to the completion of* ***one years service****, are eligible to join the BUPA Well Woman Scheme.* (BNC HP9 131)

b. *… the court had imposed a sentence of* ***one months' imprisonment*** *for the theft …* (BNC FBK 1247)

c. *… agricultural colleges offer full-time post-degree and post-HND courses, generally of* ***one-year's duration*** *…*
(BNC GUV 692)

The reader might be tempted to dismiss the spellings of N1 in (23a) and (23b) as errors, and from a prescriptivist stance that might be true. However, recall again that we are taking the sound of a construct to be its salient form and that apostrophe use is not correlated with semantics, except in a very small subset of the data. Nothing in (23a) or (23b) suggests that the writers are less than competent users of English; in fact, in our dataset, in contrast to that of Biber et al. (1999: 293), when the numeral is 'one' and N1 has an S-form, it most frequently occurs without an apostrophe, *s'* being the second most frequent form and *'s* least frequent of all. Further evidence against a common case plural analysis comes from the types of expressions exemplified in (24), where the indefinite article does not agree with the uncountable N2 but is functioning as a quantifier of N1 with the meaning 'one':

(24) a. *Andrew Pask was remanded in custody and will appear in court again in* ***a weeks time****.* (BNC K1V 427)

b. *The River Charwell burst its banks in Banbury after* ***a months rain*** *fell in just two days.* (BNC K21 3165)

To sustain an analysis of S-forms in time-measurement expressions as common-case plurals (cf. Biber et al. 1999: 293), it would be necessary to assume that, when the numeral is 'one' we have a different construction. But since cases like *one years service* and *one months' imprisonment* have the formal and semantic properties typical of the TIME-MEASUREMENT construction, we conclude that these examples are representative, and that the S-form does not therefore have common case in the construction generally.

There remain two possible analyses of the N1 S-form, both of which support the recognition of a TIME-MEASUREMENT construction. The first possibility is that it instantiates a genitive. But, as mentioned in Section 2, if the S-form is genitive then it deviates from the canonical genitive in several respects. Firstly, given a pseudo-partitive interpretation, with e.g. *six years' imprisonment* paraphrased as 'six years of imprisonment', the apostrophe appears to be attached to the 'wrong' noun. Secondly, even in cases where an alternative construal seems possible, e.g. *six years' imprisonment* – 'imprisonment of six years', these constructions deviate from canonical patterns in allowing both paraphrases and hence being ambiguous between a pseudo-partitive and a genitive reading. Not only that, but the putative genitive reading deviates markedly from the central possessive meaning of the *s*-genitive. Finally, if time-measurement expressions are genitives, then

they are atypical in being able to occur in the same noun phrase as (another) central determiner; this is what has led to them all being analysed as modifiers in the reference grammars.

The second possibility is to regard the S as neither plural nor genitive but rather as a kind of linking element. This might explain why writers are so unsure about whether to use an apostrophe in the construction and would again represent a deviation from canonical patterns. In fact, in a Construction Grammar framework, where we are only concerned with the link between surface form and meaning, it is not necessary to decide between these alternatives. We can simply recognise that the construction includes an S-form N1, with orthographic variants, and has pseudo-partitive meaning.

Another notable feature of the TIME-MEASUREMENT construction is that N1 often seems to represent the semantic head, since N2 could be omitted with little change in meaning. In some cases, such as (25), N1 is the only possible head; sometimes, as in (26), both N1 and N2 are possible heads, but omission of N1 would result in a different reading; in yet other cases, e.g. (27), N2 seems to be the head, but omission of N1 makes the expression feel semantically incomplete (at least to the authors of this chapter).

(25) a. *The course is for **four years' duration**.* (BNC B3C 1912)

b. *The course is for **four years**.*

c. **The course is for **duration**.*

(26) a. *In **ten years' time** you'll know why you were right to send off this coupon.* (BNC CFT 97)

b. *In **ten years** you'll know why you were right to send off this coupon.*

c. *In **time** you'll know why you were right to send off this coupon.*

(27) a. *Alumasc Systems are superbly engineered by a British company which has **over 40 years experience**.*
(BNC CFT 2554)

b. **Alumasc Systems are superbly engineered by a British company which has **over 40 years**.*

c. ?*Alumasc Systems are superbly engineered by a British company which has **experience**.*

The fact that N1 can sometimes function as the semantic head is one piece of evidence that, in the TIME-MEASUREMENT construction, the num-N1 element heads a phrase, namely a quantifier phrase that, in the terminology of the reference grammars, functions as the determiner of N2.

In the TIME-MEASUREMENT COMPOUND construction, exemplified by *... events in France triggered a **three-week general strike***, the bare-form of N1 does not agree in number with the numeral, except in the minority of cases where the numeral is 'one'. This lack of agreement, which constitutes a deviation from canonical patterns, is one piece of evidence for constructional status. Furthermore, since N1 does not agree in number with the numeral, num-N1 does not have the form of a phrasal construction; rather, the lack of inflection is characteristic of the modifier elements of compounds. The analysis of num-N1 as a modifier when N1 has the bare-form is further supported by the strong tendency of these expressions to be preceded by determiners which, since N1 cannot head a phrase, must belong with N2: *a general strike* not **a three week*. We even find examples like *Paul Loughlin has agreed a new **two year contract***, where the num-N1 element follows another modifier of N2, in this case *new*, underscoring the status of num-N1 as a modifier in this construction. Overall, the TIME-MEASUREMENT COMPOUND construction represents one of the lower-level (more specific) schemas (Langacker 1999; Goldberg 2002) that instantiate the abstract ENDOCENTRIC COMPOUND construction, in which a usually uninflected left-hand element modifies a right-hand head (Hilpert, 2019: 94; Booij 2009: 201).

6.2 Variation in the TIME-MEASUREMENT COMPOUND construction

Having considered the form and meaning of the central exemplars of our two constructions, i.e. those with the most frequent combinations of properties, we now turn to consider some examples where not all of these properties are present, starting with variations in the TIME-MEASUREMENT COMPOUND construction. Bare-form N1 is the most characteristic form feature of this construction, usually in combination with a preceding determiner and countable N2. However, we also find cases with bare-form N1 in which one or both of these other properties are absent. In examples with bare-form N1, countable N2, and no determiner, N2 is usually plural and represents a number of bounded entities, usually with duration num-N1: for example, 'breaks each lasting two nights' in (28a) or 'courses each lasting one day' in (28b):

(28) a. *Trusthouse Forte has **two-night breaks** in its Dublin International Hotel …* (BNC A70 2668)

b. *Noble Lowndes runs **one-day courses** for individuals whose employers do not make in-house arrangements.* (BNC CMK 171)

Singular N2s can also occur without determiners. In the TIME-MEASUREMENT COMPOUND construction, this usually happens in the context

of newspaper headlines, where the expected determiner is omitted, as in (29):

(29) a. *Pressure mounts as Bond shares slide to* ***10-year low***
(BNC A4F 319)

b. *Trapped miners crawl to safety after* ***15-hour ordeal***
(BNC AJD 731)

When an uncountable N2 occurs with bare-form N1, the uncountable noun can represent a bounded entity. This happens especially in the context of possessive noun phrases, as shown in (30). In (30a), *ten-day leave* refers to a bounded period of holiday lasting ten days, and in (30b), *21-year rule* refers to the bounded period during which a particular regime was in power.

(30) a. *Anne managed to get a few days off work during his* ***ten-day leave*** ... (BNC G16 1936)

b. *... marking the end of the* ***21-year rule*** *of President Mohammed Siyad Barre ...* (BNC HL3 689)

Whereas noun phrases with uncountable heads and no determiner must be unbounded, a definite noun phrase with an uncountable head (*his leave*, *the rule*) is ambiguous between a bounded and an unbounded reading

(Jackendoff 1996). In the TIME-MEASUREMENT COMPOUND construction, however, the meaning is constrained to the bounded reading. This constraint means that such expressions are semantically non-compositional: the meaning of the expression is not simply a combination of the independent meanings of its parts since the interpretation of N2 is constrained by the construction itself. This non-compositionality is another piece of evidence for constructional status.

When an uncountable N2 occurs with a bare-form N1 in the absence of a determiner, the construction usually refers to some kind of regular bounded event. For example, in (31a), *seven-day trading* is a kind of trading that happens on every day of the week and, in (31b), *24-hour protection* is a kind of protection that lasts for 24 hours every day:

(31) a. *... can the Minister tell us whether shops ... will have their rateable values reassessed on the basis of* ***seven-day trading****?* (BNC HHV 6511)

b. *But, he maintains,* ***24-hour protection*** *for the Ks was simply not possible.* (BNC AA1 606)

There are very strong collocational preferences in this sub-type of the TIME-MEASUREMENT COMPOUND construction; in 38% of cases num-N1 is *24-*

hour, or an orthographic variant thereof, and in a further 10% of cases it is some variant of *one-day*.

6.3 Variation in the TIME-MEASUREMENT construction

As with the TIME-MEASUREMENT COMPOUND construction, we find examples of the TIME-MEASUREMENT construction where not all central features are present. Firstly, there are examples with S-form N1 and no determiner, but countable N2. In such cases, N2 is again usually plural but the interpretation is different from that described above for plural N2 in the TIME-MEASUREMENT COMPOUND construction. When N1 has the S-form, plural N2 represents an unbounded collection of entities of which N1 represents a quantity; these types can often be paraphrased with 'worth': for example, 'three weeks worth of imports' in (32a) or 'one weeks worth of papers' in (32b).

(32) a. *Foreign currency reserves had plunged to some US$2,300 million in early March, barely enough to cover* ***three weeks' imports****.* (BNC HL6 1341)

b. *And what if someone misses* ***one weeks papers****?* (BNC J1H 3501)

However, when a singular N2 occurs with S-form N1 and no determiner, N2 is usually coerced to have an unbounded reading. This happens particularly with movement nouns, as in (33), where the time-measurement expressions do not represent bounded events ('a walk' or 'a drive'), but rather distances quantified in terms of amount of walking or driving:

(33) a. *A wide range of habitats and resources was thus available within **one or two hours' walk**.* (BNC H8U 25)

b. *… the quiet village of Mitchell, **15 minutes drive** from Truro and Newquay in Cornwall.* (BNC C9X 997)

As we saw for the TIME-MEASUREMENT COMPOUND construction, this coercion means that such expressions are semantically non-compositional, since the interpretation of N2 is constrained by the construction itself. Such non-compositional semantics is further evidence for constructional status. In fact, the TIME-MEASUREMENT construction is what is known as a GRINDING construction (Hilpert 2019: 15; Fillmore et al. 2012), in which a countable noun is coerced to an uncountable reading when it occurs in the construction (cf. the 'universal grinder'; Jackendoff 1991; Pelletier 1975). Constructional status is also supported by the tendency of this coercion to occur in time-measurement expressions with movement nouns, which might be regarded as an idiosyncratic constraint of the construction. Furthermore, the fact that

the TIME-MEASUREMENT construction constrains N2 to represent an unbounded entity or set of entities, even if N2 is formally a singular count noun, reinforces the parallel with the PSEUDO-PARTITIVE construction, where the second noun is "a bare plural count nominal or a singular mass nominal" (Falco & Zamparelli, 2019).

Finally, there are cases where an S-form N1 occurs with a preceding determiner. Payne and Huddleston (2002) suggest that this possibility is evidence that the num-N1 element is functioning as a modifier. However, their analysis depends on two assumptions, firstly that a noun phrase cannot have more than one determiner and secondly that any determiner belongs with N2 rather than N1. But in fact, we readily find examples in our data where the determiner belongs to N1, especially when N2 is uncountable. In some cases, this is evident from number marking on the determiner itself; in (34), for example, *those* cannot agree with *operation* but only with *years*:

(34) *This makes a grand total of 1,031 tons 14 cwts. for those first **12 years operation**.* (BNC H0B 563)

In other cases, number agreement is not a good diagnostic since syntactically plural measure expressions can be construed as representing single collective entities (Huddleston & Pullum 2005: 89; Keizer 2007: 122). For example, in (35), it is clear for semantic reasons that *15 seconds*

heads the subject of the subordinate clause, despite the singular verb form, since only a period of time can elapse:

(35) *As soon as the **15 seconds' rest** has elapsed, you must start the next exercise.* (BNC A0W 187)

In (35) *rest* could acceptably be omitted but *15 seconds* could not; we conclude both that num-N1 is being conceptualised as a unitary period of time, and that the definite determiner belongs with N1. There are also cases where the indefinite article combines with S-form N1 and a preceding adjective, as exemplified in (36):

(36) a. *It ended up being a great **three days music** …*
(BNC ED7 819)

b. *Not a bad **three days' work** all in all.* (BNC K4V 2462)

In these cases, the indefinite article does not agree with the uncountable N2, *music* or *work*. Rather, it agrees with the num-N1 element, *three days*, which is again being construed as a collective whole. In our time-measurement expressions, wherever a preceding determiner forms a noun phrase with num-N1, the resulting phrase represents a quantity of the

unbounded entity (or the unbounded collection of entities) represented by N2.

In some time-measurement expressions, the scope of the determiner is unclear. For example, in (37) neither *two years* nor *military service* could be omitted without impairing the coherence of the sentence and it is unclear which is best regarded as the semantic head.

(37) *... of the 33 men in the platoon, all but five are draftees whose* ***two years' military service*** *include one in Vietnam.* (BNC EE1 128)

In such cases, it is difficult to establish the intended bracketing: whether *whose* in (37) belongs with *two years* – in which case *whose two years* might be regarded as a noun phrase functioning as a quantifying determiner – or with *military service*, in which case *two years* might be regarded as a nominal modifier. However, in this example, both the anaphoric phrase *one in Vietnam*, and the plural number agreement of the verb (*include*), suggest that the intended reading is quantifying. Moreover, even if we were to conclude that *whose* has scope over *military service*, it would still be possible to regard *two years* as a post-determiner along the lines of *three* in *my three children* (Quirk et al. 1985: 261), so that the quantifier analysis could be maintained. As we have seen, the argument for analysing all S-

form time-measurement expressions as modifiers rests on the assumption that they are genitives and would therefore occupy the central determiner slot if they were to function as determiners. The same assumption also underlies the prescriptive advice to write such expressions with an apostrophe. However, despite a similarity to the *s*-genitive in phonological form, these expressions are semantically quantifying, and we therefore analyse them as occupying the same post-determiner slot as cardinal numerals and other quantifying determiners. However, as we saw with the status of the S-form, ambiguity apropos traditional grammatical categories is a feature of the TIME-MEASUREMENT construction. From a Construction Grammar perspective, the important thing is the mapping of form to function. Specifically, time-measurement expressions where N1 has the S-form and N2 represents an unbounded entity or entities have pseudo-partitive meaning.

6.4 Time-measurement phrasal compounds and ambiguous types

Despite the very strong tendency of S-form N1 to be associated with unbounded N2, there are some examples in our data where N1 has the S-form but N2 represents a bounded entity. This can happen when N2 is countable and the construction is preceded by a determiner; in other words, when the central features of the TIME-MEASUREMENT construction, apart

from S-form N1, are lacking. In such cases the coercion effect, forcing an unbounded reading of N2, does not occur, and these types do not belong to the TIME-MEASUREMENT construction. Examples are shown in (38):

(38) a. *a **10 minutes walk** down a steep drive and across a pebble beach will bring you to the town centre of Lipari.*
(BNC ECF 1598)

b. *The dissolution of the monasteries by the French in 1809 was the violent culmination of a **twenty years' campaign***
(BNC FB7 509)

In such cases, despite the form of N1, the semantics is that of the TIME-MEASUREMENT COMPOUND construction: in (38a) *10 minutes walk* represents a kind of walk with duration 10 minutes, and in (38b) *twenty years' campaign* represents a kind of campaign with duration 20 years. These examples might therefore be regarded as atypical representatives of the TIME-MEASUREMENT COMPOUND construction. However, since *10 minutes* and *twenty years* are formally noun phrases, they are also members of the PHRASAL COMPOUND construction.

Most time-measurement expressions clearly instantiate either the TIME-MEASUREMENT construction or the TIME-MEASUREMENT COMPOUND construction. However, there are some examples that are ambiguous

between the two types, usually involving a determiner whose scope is open to interpretation.

(39) a. *It is also only a pleasant **ten minutes walk** from the resort centre of Sirmione.* (BNC ECF 3681)

b. *I shall ever remember the **two or three hours conversation** I had with him that beautiful May morning he died …* (BNC ACA 666)

For example, in (39a), *a pleasant ten minutes walk* could be construed either as a quantity (a pleasant ten minutes) of walk[ing] (unbounded) or as a pleasant walk (bounded) with duration ten minutes. Similarly, in (39b), *the two or three hours conversation* could be construed either as a quantity (the two or three hours) of conversation (unbounded) or as the conversation (bounded) with duration two or three hours.

These ambiguous types highlight another parallel between time measurement expressions and *s*-genitives, where the scope of a preceding determiner also distinguishes between two constructions. Biber et al. (1999: 294) call these 'specifying genitives' and 'classifying genitives', while Quirk et al. (1985: 1335-6) refer to them as 'determinatives' and 'premodifiers', exemplified by (40a) and (40b), respectively.

(40) a. *I visited [[his old friend]'s cottage]*

b. *I visited [his [old [fisherman's cottage]]]*

In specifying genitives, which have possessive meaning, preceding determiners and modifiers belong with the genitive noun alone; for example, in (40a), *his* and *old* belong with *friend*, and the meaning is that the cottage belongs to the old friend. In contrast, classifying genitives represent a kind of the entity represented by the head noun, and any preceding determiner or modifier has scope over the whole phrase; for example, in (40b) *his* and *old* belong with *cottage*, and a fisherman's cottage is a type of cottage. Both Biber et al. (1999: 295) and Quirk et al. (1985: 1336) conclude that, whereas specifying genitive plus head noun combinations are noun phrases (e.g. *friend's cottage*), classifying genitive plus head noun combinations behave like compounds (e.g. *fisherman's cottage*). In these respects, specifying genitives and classifying genitives resemble the TIME-MEASUREMENT construction and the TIME-MEASUREMENT COMPOUND construction respectively. But, whereas the classifying genitive and the TIME-MEASUREMENT COMPOUND construction may both be kinds of endocentric compound, we reiterate that the TIME-MEASUREMENT construction differs markedly from the canonical specifying genitive both in having pseudo-partitive semantics and in functioning as a quantifier.

7. Conclusion

We can now summarise the properties of our two constructions. The TIME-MEASUREMENT COMPOUND construction is shown in (41). In has the overall form and distribution of a nominal, so that any determiner that precedes it is external to the construction and such a determiner is often required to combine with the construction to form a noun phrase. The construction itself consists of a cardinal numeral followed by a time-noun, usually with the bare-form, followed by optional modifiers and then a second noun that represents a bounded entity or entities. Semantically, num-N1 functions to modify N2 so that both in form and meaning the construction is a kind of endocentric compound.

(41) TIME-MEASUREMENT COMPOUND construction

[[Num $N1_{time}$]$_{Nmod}$ (AdjP) $N2_{bounded}$]$_{N}$

'a type of N2 with some relation to duration [Num N1]'

In addition to representing a clear pairing of form and function, the TIME-MEASUREMENT COMPOUND construction also fulfils other criteria for constructional status. It deviates from canonical patterns in lacking number agreement between the numeral and N1 and has non-compositional meaning

in that mass nouns can be constrained to bounded readings. There are collocational preferences for N2 lemma, with ‘period’ accounting for 12.3% of tokens, followed by ‘term’ (4.4%), ‘course’ (3.9%), and ‘sentence’ (2.8%). When N2 represents a regular bounded event, there are very strong collocational preferences for num-N1, with 38% of such tokens including *24-hour*, or an orthographic variant thereof, and a further 10% of cases including some variant of *one-day*. Regarding frequency, our dataset includes 11,838 tokens with the central properties of the TIME-MEASUREMENT COMPOUND construction, suggesting that the construction occurred about 118 times per million words in late twentieth century British English. This is a similar frequency to that of the word-form *July* in the corpus. Since the form-meaning pair /dʒuːˈlaɪ/ - ‘seventh month of the year’ is self-evidently entrenched in the lexicons of many English speakers, we assume that the TIME-MEASUREMENT COMPOUND construction is also sufficiently frequent to be entrenched.

Our second construction, the TIME-MEASUREMENT construction, is shown in (42). In has the overall form and distribution of a noun phrase. The determiner slot is filled by an obligatory quantifying expression which also has the form of a noun phrase and consists minimally of a cardinal numeral, followed by a time-noun ending in one of the allomorphs of S. Any determinative preceding this element is either part of the quantifying phrase or occupies a ‘central’ determiner slot with the quantifier phrase functioning

as a ‘post-determiner’ in the overall construction. There are optional modifiers and then a second noun which represents an unbounded entity or an unbounded collection of entities. Semantics-wise, the construction is quantifying, that is, num-N1 represents an amount of time in terms of which the entity represented by the second noun is measured.

(42) TIME-MEASUREMENT construction

[[(Det) Num $N1_{time}$ S]$_{NPquant}$ (AdjP) $N2_{unbounded}$]$_{NP}$

‘quantity [Num N1] of N2’

In proposing that this cluster of formal and semantic properties should be recognised as an independent construction, we are highlighting that although it resembles a genitive construction in form, it is semantically closer to the PSEUDO-PARTITIVE construction. The term ‘genitive of measure’, under which the construction is sometimes discussed as a subclass of genitive, is therefore misleading. The construction deviates semantically from the canonical pattern of the genitive construction and deviates formally from the canonical pattern of the pseudo-partitive. It has non-compositional meaning in that count nouns are coerced to unbounded readings. There are very strong lexical preferences for N2; over half of tokens with the central properties of the construction involve one of only five N2 lemmas, namely ‘time’ (21.6%), ‘imprisonment’ (11.3%), ‘service’ (8.0%), ‘notice’ (5.1%),

and ‘experience’ (4.5%). Although ‘time’ is also by far the most frequent noun lemma in the BNC generally, it occurs in the TIME-MEASUREMENT construction even more often than would be expected from its general distribution, χ^2 (1, N=100m), 90738, p < 2.2e-16. Finally, since there are 3,564 tokens in the BNC with all central properties of the TIME-MEASUREMENT construction, the same frequency as the word form *delivery* (i.e. the established form-meaning pair /dɪˈlɪvərɪ/ – ‘the action of handing something over’), we assume that the construction is sufficiently frequent to be entrenched in the minds of speakers.

References

Akaike, Hirotugu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6): 716–723.

Baayen, Harald R. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, Harald R., Davidson, Douglas J. & Bates, Douglas M. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4): 390–412.

Baayen, Harald R., Piepenbrock, Richard & Guilkers, Leon. 1995. *The CELEX Lexical Database (CD-ROM)*. Philadelphia: Linguistic Data

Consortium.

Baayen, Harald R. & Shafaei-Bajestan, Elnaz. 2019. languageR: Analyzing linguistic data: A practical introduction to statistics. R package version 1.5.0. <https://CRAN.Rproject.org/package=languageR>.

Bates, Douglas, Mächler, Martin, Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1–48.

Bauer, Laurie, Lieber, Rochelle & Plag, Ingo. 2013. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.

Bell, Melanie J. 2012.The English noun-noun construct: a morphological and syntactic object. *On-line Proceedings of the Eighth Mediterranean Morphology Meeting (MMM8) Cagliari, 14–17 September 2011*: 59–91.

Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Booij, Geert. 2009. Compounding and construction morphology. In *The Oxford Handbook of Compounding*, Rochelle Lieber & Pavol Stekauer (eds.), 201–216. Oxford: Oxford University Press.

Bybee, Joan L. 2013. Usage-based theory and exemplar representations of constructions. In *The Oxford Handbook of Construction Grammar*,

Thomas Hoffmann & Graeme Trousdale (eds.), 49–69. Oxford: Oxford University Press.

Cambridge University Press. n.d. *Cambridge dictionary online* <https://dictionary.cambridge.org/> (4 June 2021)

Croft, William 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.

Croft, William & Cruse, Alan D. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Diessel, Holger. 2019. The Gram*mar Network.* Cambridge: Cambridge University Press.

Falco, Michelangelo & Zamparelli, Roberto. 2019. Partitives and partitivity. *Glossa: A Journal of General Linguistics* 4(1): 111. DOI: http://doi.org/10.5334/gjgl.642

Fillmore, Charles J., Lee-Goldman, Russell R. & Rhodes, Russell. 2012. The framenet constructicon. In *Sign-based Construction Grammar*, Hans Christian Boas & Ivan A. Sag (eds.), 309–372. CSLI Publications/Center for the Study of Language and Information.

Goldberg, Adele E. 2002. Surface generalizations: an alternative to alternations. *Cognitive Linguistics* 13(4): 327–356.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press on Demand.

Goldberg, Adele. 2013. Constructionist approaches. In *The Oxford Handbook of Construction Grammar*, Thomas Hoffmann & Graeme Trousdale (eds.), 15–31. Oxford: Oxford University Press.

Gries, Stefan Th. & Stefanowitsch, Anatol. 2004. Extending collostructional analysis. a corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9: 97–129.

Hengeveld, Kees & Mackenzie, J. Lachlan. 2008. *Functional Discourse Grammar: A Typologically-based Theory of Language Structure*. Oxford: Oxford University Press.

Hilpert, Martin. 2019. *Construction Grammar and its Application to English 2nd edition.* Edinburgh: Edinburgh University Press.

Hoffmann, Sebastian, Evert, Stefan, Smith, Nicholas, Lee, David & Berglund-Prytz, Ylva. 2008. *Corpus Linguistics with BNCweb - A Practical Guide. English Corpus Linguistics Vol. 6*. Frankfurt am Main: Peter Lang.

Hoffmann, Thomas & Trousdale, Graeme. 2013. Construction Grammar: introduction. In *The Oxford Handbook of Construction Grammar.* Thomas Hoffmann & Graeme Trousdale (eds.), 1–13. Oxford: Oxford University Press.

Jackendoff, Ray. 1991. Parts and boundaries. *Cognition* 41: 9–45

Jackendoff, Ray. 1996. The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language & Linguistic Theory* 14(2): 305–354.

Keizer, Evelien. 2007. *The English Noun Phrase: The Nature of Linguistic Categorization.* Oxford: Oxford University Press.

Langacker, Ronald. 1995. Possession and possessive constructions. In *Language and the Cognitive Construal of the World.* John Taylor & Robert MacLaury (eds.), 51–79. Berlin: Mouton de Gruyter.

Langacker, Ronald W. 1999. *Grammar and Conceptualization*. Berlin: De Gruyter.

Lin, Jing, Hacohen, Aviya & Schaeffer, Jeannette. 2018. The interpretation of the mass-count distinction across languages and populations: introduction. *Glossa: A Journal of General Linguistics* 3(1): 70. 1–10, DOI: https://doi.org/10.5334/gjgl.638

Lyons, John. 1977. *Semantics*. 2 vols. Cambridge: Cambridge University Press.

Payne, John & Huddleston, Rodney. 2002. Nouns and noun phrases. In *The Cambridge Grammar of the English Language,* Rodney Huddleston & Geoffrey Pullum (eds.), 323–524. Cambridge: Cambridge University Press.

Pelletier, Francis Jeffry. 1975. Non-singular reference: some preliminaries.

Philosophia 5(4): 451–465

Plag, Ingo, Lohmann, Arne, Ben Hedia, Sonia & Zimmermann, Julia. 2020. An s is an ’s, or is it? plural and genitive plural are not homophonous. *Complex Words: Advances in Morphology*: 260-292.

Quirk, Randolf, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Rosenbach, Anette. 2006. Descriptive genitives in English: a case study on constructional gradience. *English Language and Linguistics* 10(1): 77–118.

Stefanowitsch, Anatol. 1998. Possession and partition: the two genitives of English. *Seminar für Englische Sprache und Kultur.*

Stefanowitsch, Anatol. 2003. Constructional semantics as a limit to grammatical variation. In *Determinants of Grammatical Variation in English,* Gunter Rohdenburg & Britta Mondorf (eds.), 55–173. Berlin: Mouton de Gruyter.

Taylor, John. 1989. Possessive genitives in English. *Linguistics* 27: 663–686

Traugott, Elizabeth Closs & Trousdale, Graeme. 2013.

Constructionalization and Constructional Changes. Oxford: Oxford University Press.

University of Oxford. 2016. *The University of Oxford Style Guide*, <https://www.ox.ac.uk/public-affairs/style-guide> (27 May 2021)

Vos, Riet. 1999. A grammar of partitive constructions. PhD Dissertation, Tilburg University, Tilburg.