

This paper has been published in *Functions of Language*. Please refer to it as follows:

Louagie, Dana & An Van linden. 2025. Noun incorporation in English: A typological perspective. *Functions of Language* 32(2). 224–278. <https://doi.org/10.1075/foL.24048.lou>

### Table of contents (hyperlinked)

1.	Introduction.....	2
2.	Data and methods.....	5
3.	Productivity of noun incorporation in English.....	6
3.1	Extensibility: Lexical expansion of the N-V pattern .....	7
3.1.1	Type and token frequencies .....	7
3.1.2	Semantic relations between IN and V .....	9
3.1.3	‘Special’ cases: Incorporation of proper nouns, plural nouns and nominal complexes .....	13
3.2	Regularity: Variety in forms.....	14
3.3	Interim conclusion .....	18
4.	Typology of noun incorporation in English .....	19
4.1	Mithun’s (1984) classification of noun incorporation .....	19
4.2	Mithun’s types I, II and IV in English .....	21
4.2.1	Type I.....	22
4.2.2	Type II.....	22
4.2.3	Type IV .....	24
4.3	What about type III? .....	28
4.4	Interim conclusion .....	30
5.	Concluding discussion .....	30
	Abbreviations .....	32
	Data availability statement .....	32
	Funding information .....	32
	Acknowledgements .....	32
	References .....	33
	Author details .....	35
	Appendix .....	36

## Noun incorporation in English: A typological perspective

Dana Louagie (University of Liège) & An Van linden (University of Liège)

This paper gives a unified account of noun incorporation in Present-Day English from a constructional and typological perspective. We first investigate how productive it really is, given the contrasting statements in the literature. Using data from the *WordBanks Online* corpus, we show that the process is remarkably productive: it is highly extensible, with high type frequency and a broad variety of possible semantic relations between the incorporated noun and its incorporating verb. It is also a regular process, with noun incorporated verbs found in all possible forms, including finite ones. Second, since noun incorporation is productive in English, we investigate how it fits in existing typologies of noun incorporation, more specifically the one proposed by Mithun (1984). We show that types I (lexical compounding) and II (case manipulation) are uncompromisingly available in English, while type III (discourse manipulation) is restricted; type IV (classificatory incorporation) is attested even if not resulting in a fully-fledged nominal classification system. We argue that these differences relate to the typological profile of English, viz. as an analytic language with overtly expressed arguments, contrary to the polysynthetic languages studied by Mithun. The availability of types II-IV incidentally further substantiates the process' productivity, since these typically develop later than type I. More generally, this study contributes to our understanding of noun incorporation and its characteristics in more analytic languages, which are understudied in this respect, and provides one of the few detailed, corpus-based studies of noun incorporation in an Indo-European language, where it is overall rare.

Keywords: noun incorporation, productivity, analogy, typology, corpus-based, nominal classification

### 1. Introduction

This paper studies noun incorporation in Present-Day English, as in the examples in (1a–d) from the *Collins WordBanks Online* corpus (WB).<sup>1</sup> Noun incorporation is “a construction in which a noun stem is combined with a verb to form a new, morphologically complex verb” (Mithun 2000: 916, referring to Sapir 1911). Thus, in (1a), the noun stem *apartment* is incorporated into the verb *hunt*, and together they make up a complex verb; the incorporated noun (IN) is semantically a patient of the verb *hunt*, which is cross-linguistically the most typical semantic relation between the IN and the verb stem (e.g. Sapir 1911; Mithun 1984, 2000; Gerdts 1998; Massam 2009); other relations are also possible, such as a locative one in (1c). A comparative example from Harakmbut is given in (2), where the patient of ‘destroy’, viz. *ak* ‘house’, is incorporated into the verb.

- (1) a. They *apartment-hunted* together, looking for a place for Elliot within a mile of their home. (WB)
- b. At school I was always in trouble, my mind was never on my work. I *clock-watched*. I would get into school at nine o'clock and spend the next six hours dreading home-time, watching the minutes tick away until eventually I had to go home. (WB)
- c. Reality TV series I'm A Celebrity Get Me Out Of Here! began on a high note when contestants were forced to *skydive* into the Australian jungle. (WB)

---

<sup>1</sup> The examples marked with (WB) were extracted from *WordBanks Online* and are reproduced with the permission of HarperCollins.

d. Rupert said: “There were some awkward scenes. There was a lot of *hand-holding* between me and Hermione.” (WB)

- (2) Harakmbut (unclassified, Peru; Van linden 2023: 470)  
*wa-mationka-eri*            *o-ak-yoŋ-me*  
NMLZ-hunt-AN            3SG.IND-house-destroy-REC.PST  
‘The hunter hut-destroyed.’

In a few places in the literature, it has been suggested that English has no significant use of noun incorporation (e.g. Mithun 1984: 847; Rosen 1989: 295; Barrie 2011: 133). Thus, for example, Mithun (1984: 847) states that “[t]he few English constructions that most closely resemble NI (e.g. *to baby-sit*, *to mountain-climb*, or *to word-process*), do not actually result from a productive compounding process, but are rather V's backformed from compound N's.” She further argues that such forms rarely occur with a full set of inflections and never without a related gerund form (Mithun 1984: 847). However, in more focused studies of English, it has been convincingly shown that noun incorporation exists and is in fact not rare at all. Hall (1956) already made a case for its increased use in English, especially since the 1900s, even if his observations were based on his own happenstance collection of items from daily life. More in-depth studies of noun incorporation in English are Rice & Prideaux (1991) and Feist (2013), both of which focus part of their discussion on the usages of noun incorporation and on explanations for why certain N-V combinations are not found. However, neither study systematically or transparently involves any quantitative analysis of corpus data. In fact, Rice & Prideaux (1991) do not seem to use corpus data at all, but at least for some data resort to speaker judgements of their own coinages.

In this study, we fill this gap by investigating noun incorporation in English from both a qualitative and a quantitative perspective, using synchronic corpus data. Following Feist (2013), we take a constructional perspective on noun incorporation, i.e. as any noun-verb compound resulting in a verb.<sup>2</sup> In other words, our approach is purely synchronic, and we disregard the etymology of individual items, some of which may result from backformation of nouns or conversion of noun-noun compounds (Hall 1956: 87; Mithun 1984: 847; Rice & Prideaux 1991: 285; Dahl 2004: 249–251). While some may argue that certain lexemes should for this reason be excluded from our study, we believe that excluding items from the start because of their etymology would keep us from thoroughly gauging the productivity of noun incorporation and uncovering the workings of analogy (see further Section 3). Moreover, if a N-V compound resulting from backformation (arguably<sup>3</sup> like *baby-sit*) synchronically behaves similarly to a N-V compound that is directly composed as such (like *star-watch* (not in OED)), we see no theoretical or empirical advantages in analysing them under a different label. The origin and diachrony of noun incorporation in English is not further discussed in this paper, but see Louagie & Van linden (2022) and Van linden & Louagie (in prep.) for a detailed study.

---

<sup>2</sup> We assume Mithun's (2000) position that noun incorporation is a morphological process with syntactic implications; we will not further be concerned with the question whether noun incorporation is in fact a morphological or syntactic process, an issue of intense debate in the literature (see e.g. Massam 2009; Mithun 2000: 924–925; Borik & Gehrke 2015: 6–9; Olthof 2020: 133–134; Haspelmath 2024: 11 for overviews).

<sup>3</sup> According to the OED, the verb *babysit* (first use 1946) is formed through compounding of *baby* + *sit*, after *babysitting* (first use 1940) and *babysitter* (first use 1937). All items start occurring around the same time, however, so the backformation analysis is perhaps less clear than it has been made out to be (see Van linden & Louagie (in prep.) for discussion).

The identification of examples as in (1) as instances of noun incorporation is further supported by the fact that they share important characteristics with other cases of noun incorporation from around the world. Feist (2013: 162–164) lists four such characteristics taken from the general literature, reproduced in (i) to (iv) below, which also apply to English:

- (i) the incorporated noun (IN)<sup>4</sup> is in a different syntactic position from its non-incorporated counterpart, and at a lower syntactic level (i.e. compound part instead of freestanding argument or adjunct);<sup>5</sup>
- (ii) the IN has a reduced morphosyntactic form or potential (e.g. losing marking for number, case, definiteness);
- (iii) the IN is semantically reduced (e.g. non-referential, with a generic interpretation);<sup>6</sup>
- (iv) noun incorporation is a marked usage.

The aim of this study is to present a unified account of noun incorporation in Present-Day English from a constructional and typological perspective. To this end, we first examine how productive noun incorporation really is in English, given the contrasting statements in the literature. The question is also interesting in light of the fact that most existing references on the topic are over 30 years old and the language may have changed in this respect; the most recent study, Feist (2013), is most positively outspoken about the productivity of noun incorporation in English. We study productivity from two perspectives: its extensibility, viz. the extent to which the N-V constructional template accepts new lexical fillers, and its regularity, viz. the range of verb forms found. Second, we investigate how noun incorporation in English compares to that in other languages, more specifically, how it fits in the four-way functional typology of noun incorporation proposed by Mithun (1984) on the basis of a wide range of languages. While earlier studies of English have briefly referred to this typology in their discussions, no systematic investigation is available. Note that Feist (2013) suggested an alternative typology for English specifically (based on the type of participant incorporated and the valency of the noun-incorporating verb), which, however, does not allow cross-linguistic comparison. Incidentally, we also disagree with some of his analyses, and will point this out at the relevant places in the paper.

This study more generally also contributes to our knowledge of noun incorporation and its characteristics in more analytic languages, which are somewhat understudied in this respect. The process of incorporation is most familiar from a range of polysynthetic languages (e.g. Mithun 1984), and Baker (1996) even argues that only polysynthetic languages truly exhibit noun incorporation (Massam 2009: 1078). Noun incorporation is, however, certainly found in analytic languages as well (e.g. Mithun 2000; see e.g. Aikhenvald (2007) for examples from various Austronesian languages). Moreover, noun incorporation is rare in Indo-European languages (Olthof 2020: 132), and not well-studied for the languages where the phenomenon

---

<sup>4</sup> Mithun (1984) stresses that it is a noun stem that is incorporated; however, convincing cases of inflected nouns or even noun phrases are found as well (see e.g. Aikhenvald [2007: 14] and Olthof [2020: 135] for examples). See further in Section 3.1.3.

<sup>5</sup> This excludes cases of conversion, such as *to hammer*, which are confusingly called instances of noun incorporation by Smirnova & Shustova (2017). This also excludes bare singular count nouns as in *He's in prison*, which are analysed as incorporated nouns by Stvan (2009) in a formalist account. Such extensions of the definition of noun incorporation are in our view not sufficiently motivated, and confuse the discussion of true cases of noun incorporation.

<sup>6</sup> While Mithun (1984: 850) maintains that INs are invariably non-referential, Baker (1988) argues that they “can be referential, i.e. they can provide antecedents for pronouns in the subsequent discourse” (Borik & Gehrke 2015: 6). However, even in formalist accounts, the latter position is no longer upheld; INs are thought to have weaker referential status (Borik & Gehrke 2015: 5).

is found. At least Frisian, German, Dutch and Danish are argued to have cases of (pseudo-)incorporation in particular constructions only (Ebert 2000; Barrie and Spreng 2009; Dähl 2004: Ch 10; Booij 2008); see also Weggelaar (1986) for an additional set of very specific cases in Dutch. More insight into an Indo-European language that clearly and quite abundantly exhibits the phenomenon is thus very welcome.

The rest of this paper is structured as follows: in Section 2, we discuss the data and methodology used in our study. Section 3 investigates the productivity of noun incorporation in English, showing that noun incorporation is an extensible and regular process, and overall very productive. The extent to which Mithun’s four-way typology can be applied to English is then discussed in Section 4. We demonstrate how all four types are found in English; type III, however, is problematic in itself, and type IV should be characterised slightly differently in order to apply to the English examples. We argue that these differences are tied to the more analytic nature of English. Section 5, finally, concludes the paper.

## 2. Data and methods

This study involves both qualitative and quantitative analysis of corpus data from the *WordBanks Online* corpus (WB). We use two datasets, approaching noun incorporation from different perspectives: one set includes data samples of 13 preselected lexemes (Dataset 1) to systematically investigate their characteristics and uses; the other casts a wider net by searching for all possible N+V combinations for specific nouns or specific verbs (Dataset 2), to study the construction more broadly.

The 13 lexemes of Dataset 1 are listed in Table 1, along with the total number of unique, relevant hits found in the corpus. The lexemes were chosen to reflect different semantic relations between the incorporated noun and the verb, such as patient (e.g. *birdwatch*), location (e.g. *skydive*), comitative (e.g. *babysit*) and instrument (e.g. *spoonfeed*), as well as different categories of incorporated nouns (animates, body parts, inanimates). For each lexeme, the dataset includes a 100-hit random sample, or an exhaustive sample when  $n < 100$ , as indicated in Table 1. The full dataset thus contains 1099 datapoints, which amounts to about one third of the exhaustive set (viz. 3004 tokens). All spelling variants are included: with a hyphen between IN and V, with IN and V in a single graphological word or with IN and V in two graphological words. Each example was annotated for inflectional (or derivational) form of the verb, transitivity, animacy of the external patient (if applicable), and type according to Mithun’s (1984) typology.

**Table 1.** Total number of tokens of the 13 pre-selected lexemes in WB, and selected sample size

lexeme	tokens	sample
babysit	1034	100
birdwatch	294	100
cherrypick	211	100
crowdsurf	42	42
giftwrap	124	100
globetrot	305	100
handhold	171	100
headhunt	181	100
heroworship	45	45

househunt	191	100
mindcontrol	12	12
skydive	292	100
spoonfeed	102	100
total	3004	1099

By contrast, Dataset 2 does not focus on preselected lexemes, but is compiled based on searches for the range of possible N+V combinations. Because an open search for any example of noun incorporation in the corpus is virtually impossible (*inter alia* because there is no tagging of part-of-speech status of compound parts), we selected individual nouns to see which verbs they can incorporate into, and vice versa. For the most part, we chose nouns and verbs among the compound parts used in the Dataset 1 (viz. *baby, bird, hand, head, hero; hunt, sit, watch*); we added two more body part nouns (viz. *brain, foot*), since we suspected these to occur in different functional types (as defined by Mithun 1984; see Section 4.1). Searches with these seven nouns and three verbs (excluding the lexemes studied in Dataset 1) resulted in a dataset of 4196 tokens in total, more than half of which are examples of the pattern *hand-V*. More details about the distribution of tokens over different patterns and the types found are discussed in Section 3.1. Each example was annotated for verb form (using only a coarse distinction between finite and non-finite, due to the large number of examples), and for type according to Mithun’s (1984) typology. An overview of Dataset 2 per type (i.e. N-V combination) is given in the Appendix.<sup>7</sup>

### 3. Productivity of noun incorporation in English

Whether or not noun incorporation is truly a productive process in English has been the subject of discussion. Mithun (1984: 847) remarked that English constructions such as *babysit* or *mountain-climb* are not yielded by productive compounding, but are rather backformations of N-N compounds. No diachronic data are given to support this claim, but four other arguments are put forward: (i) such verbs are scarce; (ii) they do not “exist without a related gerund form” (Mithun 1984: 847), (iii) they do not regularly use the full range of verb inflections, and (iv) the more versatile verbs, like *babysit*, are precisely those that do not involve the typologically typical semantic relations between the incorporated noun and the verb (viz. patient, location or instrument (see also Mithun 1984: 875)) but rather “the looser semantic relationships generally found between the constituents of nominal compounds” (Mithun 1984: 847). Points (ii)–(iii) are also stressed by Rice & Prideaux (1991) in their study of ‘object incorporation’ in English. Feist (2013), by contrast, argues that noun incorporation is productive in English, showing variation in construction types (e.g. in the relation between the incorporated noun and the verb); his claim that a variety of inflectional forms is found for at least some N-V combinations remains largely unsubstantiated, however. Note that Hall (1956) already documents several finite examples. None of the studies is backed up by any systematic analysis of English data, which makes it hard to assess the validity of their claims.

---

<sup>7</sup> Due to the nature of the searches, Dataset 2 does not include tokens where IN and V are separated by a space; from a practical perspective, it would be nearly impossible to narrow down a search to just examples exhibiting noun incorporation rather than simply an argument and verb in juxtaposition. In addition, all searches were limited to compounds tagged as verbs, which compromised recall (e.g. nominal gerunds were not retrieved), but kept the dataset manageable.

In this section, we use the two datasets introduced in Section 2 to study the productivity of noun incorporation in English from two angles. First, we investigate the ability of the N-V constructional template to take different lexical ‘fillers’ in its schematic slots, and thus form new combinations. In other words, we measure the extensibility of the pattern, which is considered the main component of productivity (Haspelmath 2002; Barðdal 2008; De Smet 2020, among many others). This will also give us insight into the semantic relations between N and V that are generally involved, and whether or not these are ‘typical’ from a typological perspective. Second, we explore the variety of forms found with noun incorporated verbs as a way to measure their regularity, a component traditionally linked to productivity (e.g. Barðdal 2008 and references therein, such as Bybee 1995, Pinker 1999). While regularity is arguably not evidence for productivity by itself but rather derivative of extensibility (Barðdal 2008), we study it here separately, since a presumed lack of inflectional variation is what kept Mithun (1984) from regarding noun incorporation as a productive process in English. We will not address the question of semantic constraints imposed on noun incorporation, since this is discussed in sufficient detail by Rice & Prideaux (1991) and especially Feist (2013).

### 3.1 Extensibility: Lexical expansion of the N-V pattern

To investigate the extensibility of noun incorporation, we surveyed the range of available N-V combinations for seven given nouns on the one hand, and for three given verbs on the other (forming Dataset 2, see Section 2). Section 3.1.1 discusses our findings in terms of type and token frequencies. Section 3.1.2. then studies the variety in semantic relations found between the incorporated noun and the incorporating verb. Section 3.1.3., finally, focuses on a few special cases where the incorporated element is not an uninflected common noun.

#### 3.1.1 Type and token frequencies

Dataset 2 contains a total of 391 unique N-V combinations. Several combinations were excluded from the dataset, because they are not transparent cases of noun incorporation. An example is *to handline*, which is a case of conversion (‘to fish with a hand-line’) that cannot be interpreted as a N-V compound in any way. A different example is *to handfast*, which was originally a N-V compound (OED); however, seeing that *to fast* is now obsolete as a verb on its own, we decided to exclude it from our dataset. By contrast, borderline examples are included when noun incorporation is a viable interpretation; these allow us to thoroughly study productivity, including the role of analogy. One such example is *to handbrake*, which can, from a synchronic perspective, be analysed in two ways: either as a conversion of the noun *handbrake* and meaning ‘to brake with the handbrake’, or as incorporation of the noun *hand* into the verb *to brake* (which exists independently), meaning ‘to brake by hand’ (compare the plethora of manner examples meaning ‘V by hand’, see below). From a diachronic perspective, we could even hypothesise that it is precisely the high productivity of noun incorporation of the pattern *hand-V* (cf. *infra*) that leads to faster conversion of items like *hand-brake*, as a N-V compound analysis is activated by analogy with similar items that were coined as cases of noun incorporation (like *hand-craft*, *hand-paint*).<sup>8</sup>

With a total of 391 unique N-V combinations for just 10 pre-defined compound parts, type frequency for noun incorporation is high. This is generally accepted to be good evidence for

---

<sup>8</sup> This would be an example of language change involving multiple source constructions (see De Smet *et al.* 2015).

productivity (e.g. Barðdal 2008; De Smet 2020 and references therein): the occurrence of many types “encourages and enables language users to abstract away from the concrete forms they directly encounter,” leading to the schematisation of the pattern, which in turn attracts new fillers (De Smet 2020: 254–255). Overall, noun incorporation is thus a productive process in English.

However, type frequency is not equally high for each pattern (i.e. of incorporations with a particular N or V). This can be seen in Table 2, which shows a break-up of the results, with the number of types as well as the total number of relevant tokens for each pattern. While the noun *hand* is incorporated into no fewer than 175 different verbs in the corpus, with a total of 2734 tokens, the noun *hero* was not found with any other verb than *worship* (see Table 1). The other patterns have varying degrees of type frequency in between these two extremes. We assume that this variation in productivity for specific compound parts can at least in part be explained by their semantics and real-world facts, such as that many “nameworthy” (Mithun 1984: 848) (and thus incorporation-worthy) actions are done ‘by hand’, while heroes are probably not so commonly referred to in everyday life. In addition, productivity seems to be an upward spiral: the more different input with one particular pattern (e.g. *hand-V*), the more new output following the same pattern (see e.g. De Smet 2020 on the effect of type frequency).

Table 2 also shows how many types have low token frequency, occurring only one, two or three times in the corpus, and how many are well-established, occurring 15 or more times in the corpus.<sup>9</sup> Across the dataset, the large majority of types have very low token frequency, while only some have a relatively high token frequency. This forms another, two-fold piece of evidence that noun incorporation in English is a productive process. On the one hand, the abundance of low-frequency items all contribute to the schematisation of the overall N-V-compound pattern, as well as of the more specific patterns with a particular N or V (e.g. *hand-V*). In fact, token frequency is generally taken to be a negative correlate for productivity, as highly frequent forms are more likely stored and processed as units, thus taking away from any potential schematisation of patterns (De Smet 2020); with only a few highly frequent lexemes in our dataset, there is nothing much detracting from the productivity of noun incorporation. On the other hand, the few high token frequency items there are, may in turn very well contribute to the productivity of the process rather than detract from it. Indeed, as argued by De Smet (2020), there is a complex interplay between different factors and “the negative effect of high token frequency [...] can be neutralized or even reversed when type frequency is high” (De Smet 2020: 270). This is because frequent exposure to a particular combination may also lead to stronger entrenchment, easier activation of the schema along with the item, and thus more productive use. A high number of occurrences for a restricted set of types may therefore also boost the schematizing effect of high type frequency (De Smet 2020), which clearly is the case in our dataset.

---

<sup>9</sup> These are more or less random cut-off points. We opted to check for ‘low frequency’ types instead of strictly hapax legomena, which is quite narrow. Nonetheless, even when we restrict the results to actual hapax legomena, the overall picture remains the same, with a total of 239 hapax legomena in the dataset.

**Table 2.** Type and token frequencies for Dataset 2 (excluding the lexemes that are in Dataset 1, viz. *babysit*, *birdwatch*, *handhold*, *headhunt*, *heroworship*, *househunt*).

Pattern	Tokens (total across types)	Unique N-V combinations (types)	Infrequent combinations (n≤3)	Well- established combinations (n≥15)
baby-V	27	15	14	0
<i>bird</i> -V	12	7	6	0
brain-V	694	39	33	2
foot-V	39	15	10	0
hand-V	2734	175	136	15
head-V	392	32	25	3
hero-V	0	0	0	0
N-hunt	108	57	53	0
N-sit	109	23	18	2
N-watch	82	29	25	2
TOTAL <sup>10</sup>	4196	391	319	25

Our dataset thus shows that all ingredients to call noun incorporation a productive process are there, perfectly matching Plag’s (2003: 54) description of productivity, as also adhered to by De Smet (2020: 272):

[U]nproductive morphological categories will be characterized by a preponderance of words with rather high frequencies and by a small number of words with low frequencies. With regard to productive processes, we expect the opposite, namely large numbers of low-frequency words and small numbers of high-frequency words.

### 3.1.2 Semantic relations between IN and V

Across Dataset 2, quite some variety is found in the semantic relations between the incorporated noun and its incorporating verb. Specifically, we find examples of all relations mentioned in Mithun’s hierarchy, supporting its validity: patient of transitive verb > patient of intransitive verbs > instruments and/or locations (1984: 875). In fact, we even find examples of other relations as well, which are represented in the more elaborate hierarchy proposed by Lehmann and Verhoeven (2005: 117): patient > theme > instrument > location/manner/time > experiencer/recipient/beneficiary > comitative > agent. This so-called ‘control’ hierarchy captures the cross-linguistic tendency that the more a participant is affected, the more readily it gets incorporated into the verb; conversely, the more a participant is in control of the event denoted by the verb, the less likely it gets incorporated. This section hence showcases the extensibility of noun incorporation in terms of the possible relations between IN and V in English.

An overview of the relations and the number of types exhibiting each relation can be found in Table 3. Not all relations are equally frequent in the dataset, but this is obviously partly due to

---

<sup>10</sup> There is one token that appeared in two searches, viz. *baby-watch*. It is included in the individual counts for *baby*-V and N-*watch* in this table, but not in the totals for types and tokens, where it is obviously only counted once.

our choice of preselected compound parts. It is for instance not surprising that the verbs ‘watch’ and ‘hunt’ almost only incorporate patients, and ‘sit’ incorporates many locations. Body parts are also naturally affected by actions and thus commonly have a patient relation with their incorporating verb;<sup>11</sup> however, they are also common in manner relations, as will be shown in this section. The frequency differences here should thus not be taken at face value.

**Table 3.** Semantic relations in Dataset 2.

Semantic relation	Number of types for 4197 tokens
Patient	187 (trans V: 184; intrans V: 3)
Manner	172
Location	18
Location-comitative	6
Comitative	5
Instrument	8
Similative	2
Beneficiary	1

By far the two most frequently found relations in the dataset are those of patient, as in (3a–c), and manner, as in (4a–b). We adopted a broad definition of patient relations, including also themes, as in (3b), where *star* is a perceived entity. Most examples involve patients of transitive verb stems (3a–b), but some involve patients of intransitive verb stems (3c). The high frequency of patient relations in the dataset is expected from a cross-linguistic perspective. A manner relation, by contrast, is perhaps more surprising; its abundance in our dataset is almost exclusively due to the wealth of lexemes of the format *hand-V*, meaning ‘to V by hand’, which is a very productive pattern in English. An exceptional example of ‘to V by foot’ is found in (4b).

(3) Patient

- a. (...) while the whole rip-roaring, empanada-selling, guitar-playing, *baby-feeding* life of the Andes goes on in the coaches behind them (...). (WB)
- b. The first time or two Hyuru and I *starwatched* together, but after that we went alone, and it was better alone. (WB)
- c. This is why a little nervousness when you take an exam is a good thing, but *hand-trembling* anxiety is not. (WB)

(4) Manner

- a. In addition to custom work in the major periods of American furniture, David also *hand-carves* architectural elements. (WB)
- b. The bottom door should have two latches, with the bottom one being *foot-operated* to save you constantly bending down to open it. (WB)

Location and instrument are also found in our data, though less commonly than patient and manner. A location relation is exemplified in (5a–c). Note the difference between (5b) and (5c), both with N-*sit*: while (5b) is purely locative, (5c) also involves a hint of a comitative

---

<sup>11</sup> McKenzie (2022: 513) even proposes the specific role of Affected Body Part in his classification of semantic roles of incorporated nouns in Kiowa and English.

relation, in analogy with lexemes such as *baby-sit* or *pet-sit* (viz. to sit ‘with’ the house and watch over it). The relation of time, which is closely linked to location, is not found in our dataset, but a dedicated search found several relevant examples, as in (6) (see Barrie & Li [2015: 177] for another example). An instrument relation can most clearly be illustrated with (7a–b); other examples are given in (7c–d) and involve body parts.<sup>12</sup>

(5) Location

- a. She had *head-mounted* lamps with big elastic bands. (WB)
- b. While I waited I hid the bat alongside the stoop between trash cans then *stairsat* watching cars swoop past. (WB)
- c. The girlfriend heard gunfire and found the man fatally wounded in the entryway of the home, which the couple were *house-sitting* for her sister, authorities said. (WB)

(6) Time

Some really got the jitters ... others, as I did, had nightmares at night. I used to wake up in the dispersal hut ... and I was *night-flying* my Hurricane [type of plane]. (WB)

(7) Instrument

- a. Last year, he said, half were killed by hunters; only one was wounded and not recovered, and that was with a rifle. Fehon (...) is so certain of the morality of *bow-hunting* that he savors every tree-hugging moment with his family. (WB)
- b. [listing items for sale] *brain-tanned* buckskin coats with buffalo collar and pewter buttons (\$500) (WB)
- c. He *hand-signaled* to the crewmen standing by to be ready to deploy the automatic nets that would capture the robot (...). (WB)
- d. A Juggler’s mad-hatter’s Hat, made by Brian in either black or brown specifically designed for manipulating and *head-catching* tricks. (WB)

The last set of semantic relations attested in English are very infrequent in our data, as well as in the world’s languages. Similitives as in (8, to handle someone as if they are a baby), for example, are not mentioned in Mithun (1984) or Lehmann and Verhoeven (2005), but McKenzie (2022: 513, 530) illustrates them for Kiowa and English. Other relations found in our dataset are beneficiary (9)<sup>13</sup> and comitative (10), which are rarely found and therefore low

---

<sup>12</sup> Body parts are not prototypical instruments, as they are inalienably possessed, but they are often treated as such in the literature (Lehmann and Verhoeven 2005: 164–167; McKenzie 2022: 530). In our analysis we distinguished between instrument and manner relations, although the difference is often very subtle. So, when *hand-signalling* (7c), the hand is used as an instrument to make signals, while *hand-writing*, by contrast, expresses the manner in which the writing is done but does not specify the writing instrument used.

<sup>13</sup> While *to proof* typically means ‘to render resistant to some force or element’ (OED) and an incorporated noun might be considered a maleficiary (as in *The bow had also been waterproofed*. (WBO)), in this case, the adaptations are made ‘so as to be safe for, resistant to, or unlikely to harm young children’ (OED), so in essence for the benefit of the baby.

on Lehmann and Verhoeven's hierarchy (2005: 117).<sup>14</sup> The latter is only found with lexemes following the pattern N-*sit*, in analogy with *baby-sit* (see Louagie & Van linden (20022) and Van linden & Louagie (in prep.) for a diachronic perspective), both with animate (10a) and inanimate INs (10b) (see also Section 4.2.3).<sup>15</sup> An agent relation is not found in our dataset, but also exists in English. An example is (11): it is the computer that controls the car.

(8) Similitive

Bobby was sobbing softly, undone perhaps by being *baby-handled*, and so reminded of his helplessness amid the fire. (WB)

(9) Beneficiary

They bought all my baby clothes, and we bought a bunch of the stuff they had to get rid of to *baby-proof* the house. (WB)

(10) Comitative

a. His obsession was so complete and his motivation so understandable that we decided to *dog-sit* for friends who were going on vacation. (WB)

b. My son, however, had other ideas. This year he is going with friends and leaving mum at home to worry and *phone-sit*. (WB)

(11) Agent

The Audis used at Silverstone are not just put onto casters, they are *computer-controlled* so that all road conditions – from dry tarmac to sheet ice – can be mimicked. (WB)

Finally, in six lexemes in the dataset, the same N-V combination may involve different semantic relations in different tokens. Four of them, incidentally all involving the pattern *hand-V*, are attested with both a manner and a patient relation between IN and V. In (12a), for example, the leather is rubbed with oil by hand (manner), while (12b) involves the action of rubbing one's hand (patient) with sanitizing lotion. A fifth lexeme, *hand-cover*, is found with a manner relation, as in (13a) (viz. covering a wall with textile by hand), and with an instrument relation, as in (13b) (viz. covering her mouth with her hand); see also footnote 11 above. The final lexeme, *brain-map*, occurs both with an internal patient relation, as in (14a), where the brain itself is mapped, and with a locational relation, as in (14b), where the herb's action is mapped *in* the brain.

(12) Manner vs. patient

a. Upon completion, Jim *hand-rubs* the leather with pure neat's-foot oil (to soften and prolong the life of the leather) (...). (WB)

b. They compared *handrubbing* with an alcohol based solution and handwashing with soap among nurses in intensive care units. (WB)

(13) Manner vs. instrument

---

<sup>14</sup> Some authors, typically formalist, have claimed that beneficiary, recipients, comitatives and agents never incorporate (e.g. Barrie and Li 2015; McKenzie 2022), but Lehmann and Verhoeven (2005: 117) cite cases in point.

<sup>15</sup> In (10b), the phone could still be argued to metonymically refer to a human patient (viz. the son, who the mother takes care of by waiting for phone calls to help him when he is in trouble).

- a. She became this sweetheart who not only cooked, but baked sourdough from a Nye family starter, who not only cleaned, but *hand-covered* a whole wall with quilted textile. (WB)
- b. Megan made her *hand-covered*, spurring giggle, watching her father and then Liz. (WB)

(14) Patient vs. location

- a. Looking for Abby Normal: Researchers have debuted a *brain-mapping* project to understand what constitutes a “normal” human brain, as opposed to an “abnormal” brain. (WB)
- b. And he is using electroencephalograms (EEGs) to “*brain-map*” the herb’s action. (WB)

### 3.1.3 ‘Special’ cases: Incorporation of proper nouns, plural nouns and nominal complexes

In this last section, we turn to cases where the incorporated element is not an uninflected common noun. These examples testify to the extensibility of the N-V pattern at a higher level of schematicity than the cases we discussed in Sections 3.1.1 and 3.1.2. We classify these special cases into three distinct types, and show that they have been observed for other languages as well.

A first type involves incorporation of a proper noun; our dataset contains two instances (15a–b). Proper nouns are not typically incorporated, since they are in principle not generic enough “to narrow the scope of V’s in a meaningful way” (Mithun 1984: 864, 866). However, the phenomenon is found in several languages around the world (see Massam 2009: 1090), and McKenzie (2022: 534) already noted it for English too. In fact, the use of incorporation indicates that the activity involving a particular, unique person is considered noteworthy by the speaker and other people in their community (e.g. the ‘circus’ referred to in (15a) is a stunt by David Blaine, who spent 44 days in a small glass box suspended above the Thames). Moreover, Mithun (1984: 861) does not preclude incorporated nouns to be associated with specific or definite reference (which is the case for the examples below); see further in Section 4.2.

- (15) a. So perhaps Londoners are more in the grip of Blaine than he is of them, in spite of all the abuse that he is taking. In that corner of central London, there is a bizarre circus geared around his every need. While I was *Blaine-watching*, there was a surreal argument between Live TV representatives and one of Blaine’s bodyguards. (WB)
- b. Then she turned to Hunter, who was demolishing a small hill of pickled herrings. “You’re *Beast-hunting* then?” she said. Hunter nodded. (WB)

A second type shows incorporation of an inflected noun, instantiated only once in our dataset (16). INs are argued to not normally inflect for categories like number, case and definiteness (see Section 1), but instances of incorporated inflected forms are found cross-linguistically nonetheless (see Olthof 2020: 134–135). The example in (16) is a borderline case, however, since it involves a plurale tantum, i.e. the plural marking is not in paradigmatic contrast with the singular.

- (16) That's a great opportunity for investors: buy on the dips (...), and sell on the spurts. That's not the advice you'd get from *fundamentals-watching* eggheads. (WB)

Finally, and most interestingly, we came across six examples (17a–f)<sup>16</sup> where an (uninflected) nominal complex is incorporated, rather than a simple noun. The first three involve a complex consisting of a noun and a classifying element: *big*, *action* and *first world war* all contribute to the denotation of subtypes of the types encoded by the head nouns *game*, *hero* and *bomb* respectively (see e.g. Halliday 1985; McGregor 1997; Davidse and Breban 2019). The fourth example is the only one involving a noun and an attributive adjective; the modified noun refers to the idiom 'born with a silver spoon in your mouth'. (Incidentally, the antonym *wooden spoon-fed* is also found on Google.) The last two examples have a complex of coordinated or disjoined nouns incorporated: in (17e), both the seizures and the mind are controlled, and in (17f), cats or dogs can be sat with. With these examples, English can be added to the list of languages allowing more complex items to be incorporated (see Aikhenvald 2007: 13–14; Olthof 2020: 135 for examples).

- (17) a. Originally introduced to Africa by *big-game-hunting* in Somaliland, he travelled southwards with his friend Dr A. E. Atkinson and entered northern BEA in 1896. (WB)
- b. For *action hero-worshipping* boys, good old Spiderman still leads the pack. (WB)
- c. The depot is the main storage facility for a *first world war bomb-hunting* team in nearby Arras, which receives thousands of calls each year to collect stray weapons. (WB)
- d. *Silver spoon-fed* [...] For those who don't have the time or talent to organise a piss up in a brewery, a new concierge service, Quintessentially, can act as your personal assistant/ contact book and life organiser. For £400 a year, it offers a 24-hour directory of where to go, how to get there and who to talk to to get the best table or access to the VIP lounge. (WB)
- e. A new medical exam showed he had no need for the *seizure and mind-controlling* drugs he had taken for so long. (WB)
- f. She charges \$12 to \$15 for *cat- or dog-sitting* and \$25 to \$40 for cleaning or party service. (WB)

In sum, the three types of special cases described above, viz. incorporation of proper names, inflected nouns or nominal complexes, are not new from a cross-linguistic perspective, but the latter two had not been described for English. Such examples remain rare, but they are nonetheless good evidence for the productivity of noun incorporation, more specifically for the extensibility of the pattern to a higher level of linguistic organisation.

### 3.2 Regularity: Variety in forms

Moving on to studying the regularity of noun incorporation, this section analyses the inflectional and derivational forms found for the 13 preselected N-V combinations in Dataset

---

<sup>16</sup> Example (16b) is part of neither Dataset 1 nor Dataset 2, but was encountered in the pre-cleaned-up dataset of *hero-worship*.

1, and adds a more cursory exploration of Dataset 2 to these findings. If noun incorporated verb forms are indeed restricted to gerunds and present participles, as suggested by Mithun (1984: 847) and Rice & Prideaux (1991: 285), this could be indicative of a more ‘noun-y’ (or at least ‘non-verb-y’) status, where such forms are then to be analysed as semi-verbalised N-N or N-Adj compounds, rather than true N-V compounds. If, however, more inflectional forms are found, this would show that these verbs are regular and fully integrated in the verbal system, supporting a hypothesis of productive N-V compound formation. In order to evaluate Mithun’s (1984) claim that only verbs with atypical semantic relationships between their compound parts have full inflectional potential (see introduction to Section 3, point (iv)), the dataset contains lexemes with ‘typical’ semantic relationships (viz. patient, instrument and location), and ones with more atypical ones (viz. comitative and simulative).

We annotated all 1099 datapoints in Dataset 1 for inflectional or derivational form at a fine-grained level, distinguishing not only the form in itself (e.g. present participle) but also the function it has in the verb phrase (VP) or noun phrase (NP). All possible forms are attested at least once in the dataset; they are illustrated in (18) below.

(18) Forms (all WB)

- a. Because I failed to watch, (...) I failed hopelessly at the *hero-worshipping*. [nominal gerund as head in a NP]
- b. Unlike past holiday seasons, many eBay sellers this year will offer *gift wrapping* services and free shipping. [nominal gerund as classifier in a NP]
- c. Bruce Wasserstein, the new chief at Lazards, has wasted no time in headhunting bankers from his old firm. [verbal gerund]<sup>17</sup>
- d. A new medical exam showed he had no need for the *seizure and mind-controlling* drugs he had taken for so long. [present participial adjective]
- e. On smaller floors buff using Liberon Drill Brush which is easily connected to any *hand held* drill. [past participial adjective]
- f. They have been spotted *house-hunting* in Cheltenham, Gloucs, but complain that every house they see is too expensive. [present participle heading a non-finite VP]
- g. These are selected from those *cherry-picked* by The MarketPlace for meeting its high standards. [past participle heading a non-finite VP]
- h. And if they’ve been *globe-trotting*, ask them, how will their jet lag improve your community? [present participle as part of a progressive VP]
- i. FORGET the rest this win has *gift wrapped* the First Division title to Falkirk with five months of the season left. [past participle as part of a perfect-tense VP]
- j. And by 4pm, all the children are home and it’s teatime — and goodness can they eat! Andrew needs to be *spoon-fed* so that can take a little while. [past participle as part of a passive VP]
- k. Women have to take responsibility. We don’t *hand-hold*. [bare infinitive]

---

<sup>17</sup> In many instances, there is no evidence for verbal or nominal status of the gerund (e.g. because they occur on their own, without objects, determiners etc.); these are annotated neutrally as ‘gerund’.

- l. If you both want to go to the wedding on 24th Oct – I’d be happy to *babysit*. [*to*-infinitive]
- m. AWAKE in the heat of the night? Slip to the fridge, let the cold air raise goosebumps on naked flesh then *spoon feed* each other a ‘love shake’ blended from ice cream or yoghurt, and tropical fruits laced with cinnamon or ginger. [imperative]
- n. While only a third of the estimated 70 million Americans who *bird-watch* every year are serious birders, the Forest Service says the total of those participating has more than tripled in 20 years. [simple present finite form]
- o. It was the city’s position that Scott Stone attended concerts before, he *crowd surfed* before, he’d been warned not to do it by his father and he chose to do it repeatedly – so we believe Mr. Stone assumed the risk when he chose to crowd surf. [simple past finite form]

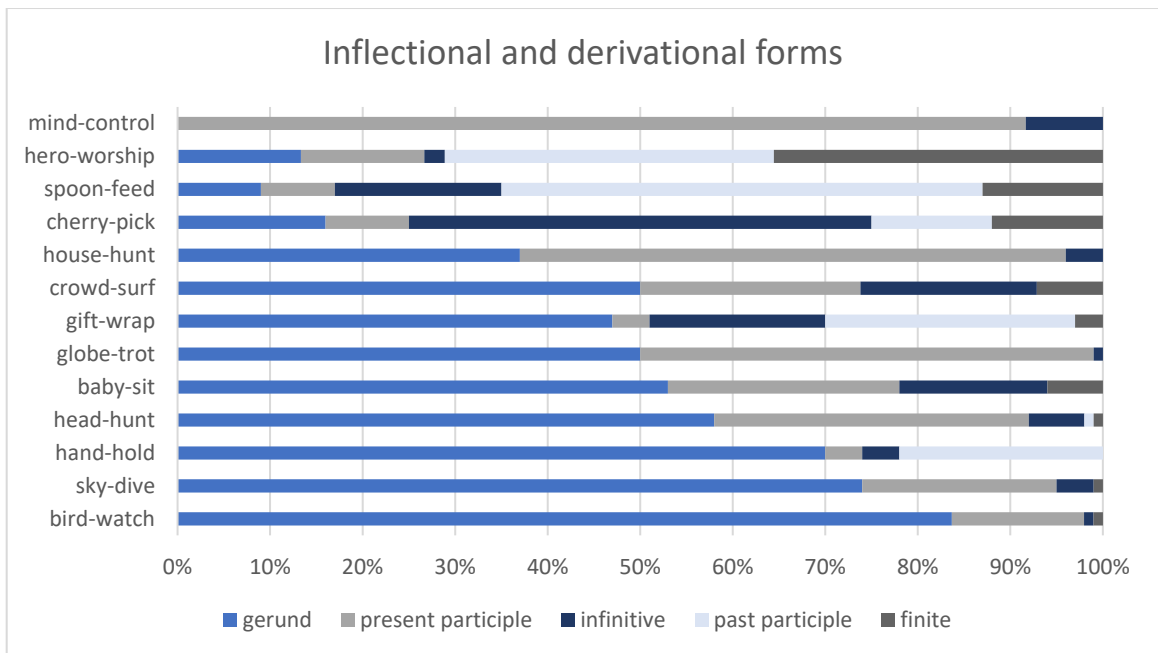
Not all forms are attested with all lexemes, however. This is partly captured in Figure 1 below, which gives an overview of the ratio of forms per lexeme (in %), with the forms merged into five categories for ease of representation: gerund, present participle, past participle, bare/*to*-infinitive and simple finite form. We find that gerunds are indeed very common, as predicted by Mithun (1984: 847), and for six lexemes make up 50% or more of the total number of tokens. Present participles are also frequently used for at least some items, like *mind-control*, *house-hunt* and *globe-trot*. By contrast, some lexemes make very little use of gerunds or present participles, but frequently occur as past participles (esp. *spoon-feed*, which is often used in the passive, as in (18j)). Simple finite forms are not highly frequent overall, but do occur for all but four lexemes.<sup>18</sup>

Figure 1 also shows that Mithun’s hypothesis concerning the inflectional potential of verbs with typical and atypical semantic relations between IN and V is only partly borne out. The four verbs with atypical relations in the dataset, viz. *hero-worship*, *baby-sit*, *gift-wrap* and *cherry-pick*,<sup>19</sup> do indeed occur with the full range of verb forms, while several verbs with a typical patient or location relation are more restricted, e.g. *bird-watch* and *globe-trot*. However, several verbs with typical relations also have full inflectional potential, e.g. *head-hunt* (patient) or *spoon-feed* (instrument), thus disproving one piece of evidence advanced by Mithun to support her claim that apparent examples of noun incorporation in English are really N-N compounds (1984: 847).

---

<sup>18</sup> For one of these four lexemes, *globe-trot*, there is a finite example in the exhaustive token set, though.

<sup>19</sup> In the last two lexemes, the IN may either bear the atypical relation of similitive to its host verb, or else the typical relation of patient.



**Figure 1.** Inflectional forms found in Dataset 1

On the one hand, the fact that the full range of possible verb forms is found across Dataset 1, including finite forms,<sup>20</sup> shows that noun incorporated verbs are treated as regular verbs, and are thus productive. On the other hand, the prevalence of participles and gerunds as well as the stark differences between individual items in the forms they use indicate that not all of these lexemes have reached full regular status (yet).

To test the validity of these findings, we also explore the verb forms found in Dataset 2, which includes a very wide range of different N-V combinations. Given the size of this dataset, we only make a coarse distinction between finite and non-finite forms. Overall, a relatively low number of examples, viz. 584 out of 4196 (14%), have a finite form. In fact, only four lexemes are responsible for about three quarters of this number (*brain-wash*, *hand-cuff*, *hand-pick*, *head-butt*), which seems to indicate that only a few items are well-established and fully regular. However, a closer look at the results shows that finite forms are actually found across 85 different N-V combinations (of the 392, see Table 2). While this number is perhaps still low, it nonetheless indicates that the full inflectional paradigm is certainly not limited to a few high-frequency items. Moreover, there is no neat correlation between overall high frequency and the occurrence of finite forms: lexemes with just one or two tokens may occur in finite form (contra Rice & Prideaux 1991: 287), while well-established lexemes may have no finite examples at all. Some finite examples of low-frequency items ( $n \leq 10$ ) are given in (19a–g). In addition, many of the finite forms are clearly no instances of backformations, contra Rice & Prideaux' (1991: 285) prediction that only the latter occur as finites. It thus seems that noun incorporated forms are regular to some extent but not fully so, confirming our findings based on Dataset 1.

(19) Finite uses of low-frequency items (all WB)

- a. Aguirre has outfitted hundreds of homes; he's worked on so many that he once *baby-proofed* the same home twice for different families.

<sup>20</sup> Note that the 'finite' category in Figure 1 excludes examples where the lexeme is part of a finite VP but not finite by itself, e.g. in a progressive form as in (18h).

- b. He sucked on the hand I bit and with the other hand he started slapping me in the face as hard as he could. It *brainrattled* me, Anne, he hit so hard and I cried I couldn't help it.
- c. I boiled so that when a man accidentally *foot-stomped* me coming out I nearly hopped up to hit him but didn't.
- d. Many of them are raised on the nursery and he *hand-selects* the seeds – other more specialist hybrid seed is bought in bulk.
- e. When night came I rose refreshed and eyed my pillow where I *headlaid*.
- f. New York markets were uneven, with blue chips positive as investors *bargain-hunted* after Wednesday's selloff, while techs finished lower on mixed earnings reports.
- g. This afternoon we went to Civics class and Ms Boardman had the TV on and we *news-watched* some more.

Two final examples may serve to illustrate the regularity of (at least some) noun incorporated verbs. First, we found a single example in the corpus which uses the non-standard past participle *baby-sitted* instead of *baby-sat* (20); Google searches yield many more such examples (and this form was already noted by Hall (1956: 87)). This example is in line with non-standard occurrences of the simple verb, where language users use a past form in *-ed* in analogy with regular verbs. Second, example (21) shows how the passive participle of a noun-incorporated verb may be prefixed with *un-*, just like passive participles of morphologically simple verbs can be, to encode that the action has not been carried out (OED, s.v. *un-*). In this case, the incorporated noun occurs between the prefix *un-* and the root *sit*.

- (20) Recently, his ex-wife had had a baby boy and on Saturday evening Dena had *baby-sitted* while Jean and her husband went to the cinema. (WB)
- (21) ...when they go to the bad, take another drink, smoke an illicit joint, leave the child *un-babysat*, leap at the genitals of another sex, why then they are Angel. (WB)

### 3.3 Interim conclusion

In this section, we have shown that noun incorporation is both an extensible and a regular process in English, using two sets of corpus data. First, the extensibility of noun incorporation can be seen both in the wide range of fillers that the N-V schema can take (viz. there is high type frequency) and in the varied semantic relations that are possible between IN and V. We even found a few examples that further stretch the N-V schema, allowing nominal complexes to incorporate into a verb. Second, both datasets contained examples of any possible verb form (including finite forms, both for well-established and low-frequency items), showing the regularity of noun incorporation. We may thus conclude that noun incorporation is productive in English.

Some nuance is necessary, however. While noun incorporation is overall a productive process, productivity is not equally strong across all 'subtypes' or individual lexemes. First, there are a few very high-frequency items, such as *baby-sit*, *hand-cuff* and *brain-wash*, which could be argued to be autonomously stored rather than being 'active' incorporations; they also skew the results slightly in that they for example use the full inflectional paradigm and are often found in finite form. We have seen, however, that the presence of some items with high token frequency need not detract from the productivity of the process in general (De Smet 2020).

Second, some nouns or verbs more easily enter into a noun incorporation construction than others. This is not surprising, and can at least partly be explained by their semantics in combination with the types of actions regularly referred to by noun incorporated verbs (viz. noteworthy ones, see Section 4.1). Some further factors that constrain the productivity of noun incorporation on the individual lexical level are discussed in more detail by Feist (2016: 179–182); see also Olthof *et al.* (2021) for a cross-linguistic perspective on verb-based restrictions. Finally, lexemes differ with respect to the variety in inflectional forms they show. All these factors show that some forms are more productive or more likely to be coined than others (e.g. with patient relations), but this does not detract from the overall productivity of noun incorporation.

#### 4. Typology of noun incorporation in English

The finding that noun incorporation is a productive process in English motivates our second research question, viz. how it fits in cross-linguistic typologies of the phenomenon. We use Mithun’s (1984) seminal typology of noun incorporation as a starting point (Section 4.1) and investigate to what extent it can be applied to the English data (Sections 4.2 and 4.3). Our findings will ultimately also feed back into our earlier question regarding the productivity of the phenomenon, with the availability of different types of incorporation serving as an index of productivity (see Section 5).

##### 4.1 Mithun’s (1984) classification of noun incorporation

Using data from a diverse set of more than 100 languages, Mithun (1984) proposes a four-way typology of noun incorporation based on its functions. In this section, we briefly summarise these four types. Type I constitutes basic lexical compounding used to refer to “name-worthy” activities (i.e. recognisable or institutionalised ones), as in (22a). Type II also involves lexical compounding for name-worthy activities, but here the use of noun incorporation manipulates the case relations in the clause: an oblique argument is ‘advanced’ into the case position that is ‘vacated’ by the incorporated noun. This type frequently involves the incorporation of body part terms, with the possessor of the body part then functioning as a main argument of the verb. An example is found in (22c), which can be compared to the non-incorporated alternative in (22b). In (22b), the object pronoun on the verb *-yos-* indexes the face, while in (22c), ‘face’ is incorporated, leaving the object position on the verb open for the possessor of the face, *-s-* ‘him’. An example of type II incorporation *not* involving body parts is given in (23): here, the incorporation of the patient ‘tree’ allows a locational phrase (‘[in] my cornfield’) to function as direct object of the verb, instead of as adjunct marked by a preposition.

(22) Tupinambá Tupi (Guaraní; Rodrigues ms., cit. in Mithun 1984: 856–857)

- a. *a-pisá-eyti'k*  
I-fishnet-throw  
'I net-throw.'
- b. *s-oβá*                      *a-yos-éy*  
his-face                      I-it-wash  
'I washed his face.'
- c. *a-s-oβá-éy*

I-him-face-wash  
 ‘I face-washed him.’

(23) Yucatec Mayan (Mayan; Bricker 1978, cit. in Mithun 1984: 858)

*k-in-č’ak-č’e’-t-ik*                      *in-kool*  
 NCOMP-I-chop-tree-TR-IPFV      my-cornfield  
 ‘I clear my cornfield.’ (lit.: I tree-chop my cornfield)

Type III noun incorporation functions to manipulate discourse structure, in that it backgrounds known or incidental information in a stretch of discourse. In (24), for instance, a referent is introduced with an external noun by speaker A, while it is incorporated in speaker B’s reply because it is old information. Strictly speaking, the incorporated noun is non-referential, but it serves an anaphoric tracking function here. More generally, in type III, INs may even be associated with specific and definite reference in their tracking function (Mithun 1984: 861).

(24) Huauhtla Nahuatl (Uto-Aztecans; Merlan 1976, cit. in Mithun 1984: 860–861)

A:     *askeman*            *ti-’kwa*            *nakatl.*  
          never                you-it-eat            meat  
          ‘You never eat meat.’  
 B:     *na’*            *ipanima*            *ni-naka-kwa.*  
          I                always                I-meat-eat  
          ‘I eat it (meat) all the time.’

In type IV, finally, noun incorporation is classificatory. A general noun stem, incorporated into the verb, may “be accompanied by a more specific external NP which identifies the argument implied by the IN” (Mithun 1984: 863). This may then result in a classificatory system, where nouns are classified according to the incorporated noun used in a construction with them. An example is given in (25), where the incorporated noun - *ič’á-* ‘eye’ classifies the external noun phrases *kassi* ‘bead’ (25a) and *ka’ás* ‘plum’ (25b) as small, round objects.

(25) Caddo (Caddoan; Chafe 1977, cit. in Mithun 1984: 865)

a.     *kassi’ háh-’ič’á-sswi’-sa’*  
          bead    PROG-eye-string-PROG  
          ‘She is stringing beads.’  
 b.     *ka’ás háh-’ič’ah-’i’-sa’*  
          plum    PROG-eye-grow-PROG  
          ‘Plums are growing.’

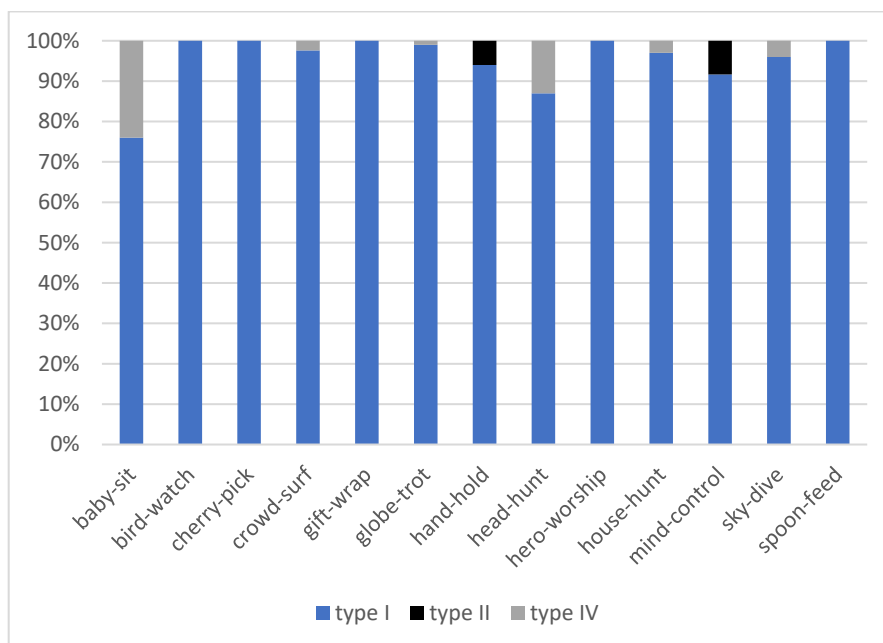
Mithun (1984: 874–877) further argues that the four types also form an implicational hierarchy: type I > II > III > IV. Thus, from a synchronic perspective, a language with type IV noun incorporation will also have all the other types, while a language with type II will also have type I (but not necessarily types III and IV). From a diachronic perspective, languages develop the types in the same order: type I first and type IV last.

In the rest of Section 4, we investigate to what extent Mithun’s typology can be applied to English synchronic data, and whether the implicational hierarchy also holds for English from a synchronic perspective. To this end, we annotated our two datasets for the four types summarised above; we discuss the results in Section 4.2. In applying the typology, we found that type III is not well-delineated in Mithun’s work, as it is the only type that is not clearly structurally distinguished from the others: while type II differs from type I in case role assignment, and type IV from the others in that the incorporated noun is accompanied by an

external noun phrase,<sup>21</sup> type III has no such clear formal/structural characteristics that set it apart from the others (cf. also Massam 2009: 1079–1080). Therefore, we did not code our data for it. However, we conducted a smaller pilot study on type III separately, which is discussed in Section 4.3.

#### 4.2 Mithun’s types I, II and IV in English

An overview of the distribution of types I, II and IV in Dataset 1 can be found in Figure 2; as explained above, type III is taken out of consideration for now. Type I is by far the most common type in the data: it is found for all 13 lexemes in the dataset, and is the most common use for each lexeme individually. Type II occurs with two lexemes; type IV with six lexemes (and a further three which constitute less prototypical instantiations and are therefore not represented in Figure 2; but see Section 4.2.3).



**Figure 2.** Mithun’s typology applied to Dataset 1.

The advantage of Dataset 1 is that it allows us to consistently compare similarly-sized samples of tokens. However, the results are obviously also influenced by the lexemes we selected. For example, type II often occurs with incorporated body parts cross-linguistically (Mithun 1984: 860), so with only a few such items in our selected set, it is to be expected that type II is infrequent in our data. To partly remedy this, we also annotated Dataset 2 for Mithun’s typology. This dataset is much wider in scope: while still based on pre-determined compound parts, it contains almost 400 different N+V combinations. The results, captured in the Appendix, largely confirm those of Dataset 1: incorporation in English is overwhelmingly of type I, while types II and IV are found in a large enough number of examples to ascertain

<sup>21</sup> Although Mithun (1984: 863) states that with classificatory NI the N-V combination “*can* be accompanied by a more specific external NP” (emphasis ours), we took this as a prerequisite in our analysis.

that their occurrence was not a result of cherry-picking of lexemes of any kind in Dataset 1. In the following subsections, we discuss each of the three types in turn.

First, however, we briefly explain how we applied the recognition criteria for Mithun's types to cases where the lexemes studied do not serve as (parts of) finite verb forms. Indeed, this study extends Mithun's dataset to include forms that serve (i) as heads of non-finite clauses, (ii) as participial adjectives, and (iii) as nominal gerunds. The first set involves a straightforward application of the same recognition criteria as with finite verb forms; see e.g. the discussion of example (28b) in Section 4.2.2. Secondly, we argue that participial adjectives, while not having an argument structure themselves, are comparable to their predicative counterparts. For instance, we analyse *brain-damaged patients* as equivalent to *the patients are brain-damaged*, and thus as involving type II incorporation; in both cases there is a possessive relation between the IN *brain* and the noun *patients*. Thirdly, the analysis for nominal gerunds differs according to their more specific function. A nominal gerund with a classifying (modifier) function within the noun phrase, as in (25e), has no argument structure at all, and can thus in principle only involve type I incorporation. Nominal gerunds functioning as head of a noun phrase can be analysed in a straightforward manner, since their arguments can be realised periphrastically (as in *the hero-worshipping of older boys*).

#### 4.2.1 Type I

Type I, or basic lexical compounding without any of the structural characteristics found for types II and IV, is widespread in English, as shown above in Figure 2 and the Appendix. Examples from both datasets are given in (26).

- (26) a. For now, though, he will *globe trot*, following his game. (WB)
- b. [...] it would be impossible to *hand hold* the camera and keep it still during the interview which would run for well over an hour. (WB)
- c. It's difficult to *people-watch* when you've been on telly and people know your face, so I put on a silly hat and a pair of glasses. (WB)
- d. The pack, greatcoat and satchel were heavy, the lanes were *foot-clogging* with mud and dung, and he knew he must find a lair soon, so he twisted into a narrow passage [...]. (WB)
- e. A major country ball promised to be an excellent venue at which to further polish her, not to put too fine a point on it, *husband-hunting* skills. (WB)
- f. 'I guess you've received tons of fan mail from *hero-worshipping* dames all over the country,' (...). (WB)

#### 4.2.2 Type II

Type II, noun incorporation for manipulating case relations, is found in English, confirming earlier findings by Rice & Prideaux (1991: 285) and Feist (2013: 170). It is much less common than type I in both our datasets. Dataset 1 has seven examples in total, involving two lexical items, viz. *hand-hold* and *mind-control*. It is not a coincidence that both these items involve the incorporation of a body part term: body parts are inherently 'possessed' and actions happening to body parts naturally affect their possessors, which in this construction

function as direct objects syntactically. This is illustrated in (27) for *mind-control*: the (metonymic) possessor of the minds that are controlled (viz. *the whole country*) is placed in direct object position, a position that is left open with the incorporation of the patient *mind*. The incorporation of the body part term thus allows a language user to highlight that the entire person is affected, and not just their body part (cf. also Mithun 1984: 858).<sup>22</sup>

(27) [He] is secretly planning a breathtaking stunt that will involve the entire population. “I want to go on television and *mind-control* the whole country,” he revealed. (WB)

An example with *hand-hold* is given in (28): again, the possessor of the hand which is held is ‘promoted’ to direct object position, which is vacant due to the incorporation of the semantic patient *hand*. The non-incorporated equivalent of these constructions is ‘to hold her hand’.<sup>23</sup> Compare this type II example with example (26b) above, also with *hand-hold* but showing type I noun incorporation. There, the non-incorporated equivalent is ‘to hold the camera in your hand’, i.e. the incorporated noun is in a locative relation to the verb stem, not in a patient relation as in example (28) below (see also Feist 2013: 170).

(28) Mama’s sleepless nearly every night now and I sometimes sit with her hours long *handholding* her and saying “it’s all right, Mama, it’s all right”. (WB)

Dataset 2 further supports the availability of type II noun incorporation in English with a relatively large set of examples, (unsurprisingly) almost all involving body-part incorporation as well. Examples using 15 different N-V compounds are found, including hundreds of instances with *hand-cuff* and with *brain-wash*. Some straightforward examples are given in (29), where the non-incorporated equivalent is always ‘to V someone’s feet/hand/head/brain’; (19b) above is another example. Note that example (29b) is passive but can still be argued to involve type II incorporation: it is the prisoner’s hands that are tied, and incorporation allows for the possessor of the body part to be used in an argument function. Note that instances may also be figurative, as in (29e);<sup>24</sup> the literal or figurative interpretation of a compound does not have any impact on the analysis in terms of Mithun’s types.

- (29) a. High fashion shoes today provide a form of torture unseen since the Chinese gave up *foot-binding* girl children by bending their toes underneath the foot. (WB)
- b. The prisoners, still *hand-tied*, went next, packed into three trucks. (WB)
- c. After seeing the disgusting thing the little scummer had scrawled on his chest, Mr. Cutter *head-bashed* the bastard with a rock, then slipped away unnoticed. (WB)
- d. To emphasize the point, the actor rapped Fang on the head and very nearly *brain-damaged* him! (WB)

---

<sup>22</sup> As expected, the incorporated noun in this example is not specified for number even though multiple minds are involved in the action.

<sup>23</sup> One might also argue that the non-incorporated equivalent is an external possession construction, ‘to hold someone by the hand’, so that the incorporated construction would involve incorporation of location rather than patient. However, there is no good reason to favour this analysis over one involving an adnominal possession construction. Moreover, there are several examples where an external possession construction would not be possible at all, as in (27) above and (29) below. We hence opted for the most parsimonious approach, analysing all constructions involving a part-whole relation between the incorporated noun and the direct object along the same lines.

<sup>24</sup> In our dataset the lexemes with figurative meanings are always still compositional in meaning, even though non-compositionality is common for noun incorporation cross-linguistically (e.g. Aikhenvald 2007: 16).

- e. But he is only interested in using their music to hide subliminal recorded messages that *brainwash* American teenagers into buying useless fashion items. (WB)

There are a few examples where multiple analyses are possible, either as type II with case role manipulation, or simply as type I without such manipulations. In (30) for example, it is unclear whether the incorporated noun *head* is ‘possessed’ by the patient undergoing the action (‘I smashed his head’), as in the previous examples, or by the actor doing the action, thus involving an instrument relation between the incorporated noun and the verb (‘I smashed him using my head’). The former analysis would imply there is type II incorporation, the latter type I. Compare this example to (29c) above, where an analysis with the head as instrument is not very likely: it is clear from the context that the patient’s head is bashed with a rock.

- (30) (...) all of this in the moment before I *head-smashed* him on his damaged nose and scraped my boot down his shin and ground it into his instep. (WB)

Example (31) does not involve body parts but arguably another type of possessive relation between the incorporated noun and the direct object of the verb (viz. ‘the scoreboard of Michael Jordan’s Wizards’). Alternatively, the IN in this example can be analysed as having a locational relation to the verb (‘watch Michael Jordan’s Wizards on the scoreboard’), in which case there is no manipulation of case relations and thus simply type I incorporation. The latter analysis is perhaps more plausible from a semantic perspective.<sup>25</sup>

- (31) Chaney admits he’s been *scoreboard-watching* Michael Jordan’s Wizards. (WB)

Finally, an exceptional example of type II is found in (32), which does not involve a possessive relation between incorporated noun and object noun phrase as in the previous examples. Rather, the incorporation of the patient (*coon* for *raccoon*) allows a locational phrase (*[on] Dunbar’s farm*) to function as direct object (much like in the Yucatec Mayan example (23) discussed in Section 4.1). This is the only example of this kind in our dataset, but another example of a locational phrase being ‘advanced’ to direct object status is offered by Feist (2016: 230), repeated in (33).

- (32) Bud told a story about how he and Snap LeClair *coonhunted* Dunbar’s farm on summer nights. (WB)
- (33) “He was awakened by police later in the morning when they *door-knocked* the street as part of their investigation.” (Daily newspaper report, cited in Feist 2016: 230)

### 4.2.3 Type IV

Type IV, classificatory incorporation, is also modestly frequently found; it is especially common with *baby-sit* in Dataset 1 (cf. Figure 2 above), but also occurs with several other lexemes in both Dataset 1 and Dataset 2. In this section, we first make our case for type IV noun incorporation in English, and illustrate this with various examples. We also discuss some differences between type IV incorporation in English and in the polysynthetic languages

---

<sup>25</sup> For examples with multiple possible analyses, like the ones in (30)–(31), the most conservative analysis is included in the counts in Table 3 and the Supplementary file, viz. as type I and involving instrumental (30) and locational (31) semantic relations between IN and V.

discussed by Mithun (1984). Finally, we discuss several cases of non-prototypical type IV noun incorporation (which were not ranged with type IV in Figure 2); such cases have gone unnoticed in the literature so far.

Contra Feist (2013: 137) but following Rice & Prideaux (1991: 285),<sup>26</sup> we argue that examples like (34a–c) instantiate type IV noun incorporation. In (34a), the external noun phrase *the grandchildren* identifies the argument implied by the incorporated noun *baby* (cf. Mithun 1984: 863), and *baby* semantically classifies the referent of *the grandchildren*. This analysis also holds for examples like (34b–c), where the direct object is an adult or inanimate entity. Such uses show that the incorporated noun *baby* has widened its semantic scope, its meaning being generalised to ‘entity in need of care’ with any restrictions regarding age or animacy lifted. We will come back to example (34b) below.

- (34) a. Sadly, I’ve missed a few matches recently *babysitting* the grandchildren. (WB)  
 b. O’Meara [...] said: “I *baby-sat* Tiger Woods and I told Tiger it’s his turn to babysit Sergio [García]. He’s an incredible talent but he is still only 19.” (WB)  
 c. Cut to Florida’s Keys and Rick Broca, a retired cop *babysitting* the yacht of his boss. (WB)

Some examples of type IV noun incorporation with other lexical items are given in (35a–f) (from both datasets), going from most to least clearly having a classificatory function. Thus, example (35a) is semantically and structurally very close to the examples in (34) above. In (35b), the classificatory relation between the incorporated noun *people* and part of the direct object, *the Catalan families*, is clear; the inclusion of *old-timers* under the same category is somewhat more surprising. Similarly, in (35c–d), *a luxury home* is a type of *house* you can hunt for, and *bankers (from his old firm)* are a type of people (with the IN *head* as pars pro toto). Finally, (35e–f) are perhaps less typical instances of type IV, where the incorporated noun expresses a location and an external locational phrase identifies this location more precisely: *an endless sea of strangers* is the crowd on which someone surfs, and *the edge of space* is a more precise location in the sky from which someone dives. The latter examples also show that there is no requirement for the classified item to function as direct object; it can also be expressed by a prepositional complement (as in (35c)) or an optional adjunct (as in (35e–f)).

- (35) a. He could relieve Gaby and Charley of *baby-watching* Steve, and perhaps do some clue-hunting in the flat. (WB)  
 b. Can Culleretes is very popular (and not full of tourists), but you can usually squeeze in for lunch, and **people-watch** the Catalan families and old-timers catching up over a glass or two. (WB)  
 c. Now they and their children [...] are **househunting** for a luxury home. (WB)  
 d. Bruce Wasserstein [...] has wasted no time in **headhunting** bankers from his old firm. (WB)  
 e. Her ideal was to have no context at all, only weightlessly to **crowd-surf** on an endless sea of strangers who would hold, fondle, and pass along every facet of her glamorous existence. (WB)

---

<sup>26</sup> To be precise, they did not refer to Mithun’s type IV, but mention the incorporated noun “serv[ing] a classificatory function vis-à-vis the independent direct object” (Rice & Prideaux 1991: 285), following Rosen (1989).

- f. It centres around his bid to claim a unique world record by **skydiving** from the edge of space. (WB)

There are at least two obvious differences between type IV incorporation in English and in the polysynthetic languages discussed by Mithun (1984). First, there are clear lexical restrictions in English: it seems that not any noun can (easily) incorporate into any verb. Hence, type IV noun incorporation cannot be used to keep a topic straight over a stretch of text with different verbs, as it does in the languages discussed by Mithun (1984); the restricted presence of type III also relates to this, as further discussed in Section 4.3.

Second, cases like (34) and (35) above do not result in a broad system of nominal classification in English, unlike in some other languages with type IV, where nouns are classified according to the general noun that is incorporated into the verbs they are arguments of (Mithun 1984: 863). However, through extensibility (see Section 3.1) and analogy, more nouns are incorporated into verbs like *sit* and *hunt*, which in a way results in a miniature and incipient system of classification. This is most clearly the case with *-sit*: in analogy with *baby-sit*, we find examples of *pet-sit*, *dog-sit*, *cat-sit*, *house-sit*, *flat-sit*, *villa-sit*, etc. Some, but not all of these, also occur in a type IV construction in the corpus. In (36a), for example, the general noun *pet* is incorporated into the verb, classifying the direct object *dog*, *cat* or *hamster*. Interestingly, this alternates both with examples like (36b) (not in our datasets but via Google), where the incorporated noun *baby* can also classify non-human referents, as discussed earlier, and with examples like (36c, repeated from 17e), where a more specific noun is incorporated into the verb in a type I construction.

- (36) a. Let a friend or neighbor *pet-sit* your dog, cat or hamster during showings. (WB)  
 b. How do I tell someone that I will not *babysit* their dog when they bring it over? (Google)  
 c. She charges \$12 to \$15 for *cat- or dog-sitting* and \$25 to \$40 for cleaning or party service. (WB)

Similarly, houses and other locations can be ‘sat (at/with)’ and cared for. Thus, (37a–b) exemplify two types of dwellings classified by the incorporated noun *house*; the widened semantic scope of *house* is evidenced by the occurrence of *an apartment* (which is not strictly speaking a house) as direct object in (37b). Compare these examples with (37c–d), where a more specific noun is incorporated instead. Interestingly, these still occur in a type IV construction: in (37c), *her apartment* identifies the referent implied by *flat*, as does the prepositional phrase *at I Tatti* the incorporated noun *villa* in (37d) (*I Tatti* being the name of a particular villa). Example (37e) presents a final illustration of a type IV construction with *N-sit*, now with a more general locational noun *site* incorporated and the buildings of the site being constructed as the direct object of the verb. This type of variation found in the dataset shows that there is no robust classification system in English as found in languages like Caddo and Gunwinygu (Mithun 1984), but rather an incipient, flexible and even ad-hoc classificatory system employed in a way that fits the English typological profile.

- (37) a. Son Simon, 28, was supposed to have been *house-sitting* the semi in Newbury, Berks. (WB)  
 b. Mack and Stan are spending the summer in New Orleans, *housesitting* an apartment just outside the French Quarter. (WB)  
 c. It was winter, almost a year after Susan’s death, and I had been *flat-sitting* her apartment while it was on the market. (WB)

- d. In fact, while they were *villa-sitting* at I Tatti for the Berensons in the winter of the same year, he was still dreaming about living part-time in Cambridge (...). (WB)
- e. One squatter said: “We are a collective of artists, healers and musicians who *site-sit* empty buildings to restore and rejuvenate them.” (WB)

The last issue to be discussed in this section concerns some less prototypical instantiations of type IV incorporation in English, which have gone unnoticed in the literature (as far as we are aware). In both datasets, there are some cases where there is no straightforward classificatory relation between the incorporated noun and an external noun phrase; rather, the classificatory relation is non-inherent, attitudinal and often comes with an additional hypothetical or resultative layer. In addition, the meaning is figurative in many examples. Perhaps unsurprisingly, these cases nearly all involve a similitive relation between the incorporated noun and the verb (see Section 3.1.2). For instance, in (38), Stalin is worshipped as a hero; that is, he is not inherently a member of the class of heroes, but he is classified as one by the agent of the clause. Note that the same analysis arguably also applies to (34b): Tiger Woods is classified as a baby or entity in need of care by the speaker; other speakers might not concur with this classification. This contrasts with the classificatory relation in examples like (37a) above: a semi *is* a type of house; the classification relies on features inherent to the referent and does not depend on the speaker’s or agent’s attitudes.

(38) He *hero-worshipped* Stalin and despised the likes of Gorbachev and Yeltsin. (WB)

Similarly, in (39a), the relation between the incorporated noun *witch* and the external NP associated with it, viz. *I*, is hypothetical, not inherent: the person is hunted as if she were a witch, i.e. she is subject to the same treatment as witches were in medieval times. The same is true for all examples with *cherry-pick* that have an external associated noun phrase: in (39b), for example, a slice of data is picked *as if* it were a cherry (i.e. carefully). Finally, with *gift-wrap*<sup>27</sup> the classificatory relation is imposed by the giver when the lexeme is used in its literal sense – the I-person considers the earrings to be a gift in (40a) –, and the relation even gets an additional hypothetical or resultative layer when the lexeme is used figuratively. Thus, in (40b), the burger is wrapped as if it were a gift (hypothetical); in (40c), Erin is wrapped so that she becomes a gift (resultative). Finally, (40d) shows a hypothetical and fully figurative use without any real or analogous wrappings, as highlighted by the adverb *virtually*.<sup>28</sup>

- (39) a. Her defense is being handled through court-appointed attorneys. “I’ve been *witch-hunted*. Pure, unmitigated witch-hunted,” said Tarnavsky (...). (WB)
- b. Others will *cherry-pick* a narrow slice of data for publication while consigning the rest to the file drawer. (WB)
- (40) a. I had them *gift wrap* the earrings, (...). (WB)
- b. The big poppyseed bun *gift wrapped* the burger, and the American cheese (my choice) quite nicely. (WB)

<sup>27</sup> Note that Rice & Prideaux (1991: 285) also analyse such instances as having a classificatory function.

<sup>28</sup> Incidentally, this example set once again shows that there is no one-to-one correlation between Mithun’s functional types and the semantic relation between IN and V: example (40a) involves a patient relation, viz. ‘to wrap a gift (which are the earrings)’, while examples (40b–d) involve a similitive relation, viz. ‘to wrap X like a gift’. A similar observation was made for ‘classic’ type IV incorporation (i.e. involving inherent classification) above, where we saw examples involving incorporated patients, comitatives or locations (34)–(35).

- c. The door to the bathroom opened. Erin was neatly *giftwrapped* in a towel, her freshly washed hair piled high on her head. (WB)
- d. Defensive blunders virtually *giftwrapped* newly-promoted Norwich the goals which earned Nigel Worthington's side a battling point. (WB)

In all these examples, we argue, the classificatory relation is non-inherent, contrary to the more typical classificatory examples discussed previously. If we then broaden the definition of type IV noun incorporation to include such non-inherent relations, the proportion of type IV in Dataset 1 is much larger than presented in Figure 2 above, and would encompass 56% of instances with *gift-wrap*, 82.2% with *hero-worship* and 85% with *cherry-pick*.

### 4.3 What about type III?

Type III noun incorporation has no structural differences with the other types (Mithun 1984: 862), and is therefore quite difficult to apply to new corpus data. We did not include it in our first round of annotations or in the results reported on in the previous section. This section more thoroughly investigates whether or not type III is available in English, based on a detailed pilot study of two samples in Dataset 1, viz. those of *bird-watch* and *baby-sit*.

Lacking a structural distinction, the only distinctive feature of type III is its function: it backgrounds known or incidental information in parts of the discourse (Mithun 1984: 859), even while remaining non-referential (Mithun 1984: 866). The difference with the other types is admitted to be “subtle”: “Type I serves to reduce [the IN’s] salience within the V, Type II within the clause, and Type III within a particular portion of the discourse” (Mithun 1984: 862). This is rather difficult to apply in practice. How to tell the extent to which the IN is backgrounded? The recognition criterion of the IN representing known or incidental information is also less distinctive than it seems at first sight: one could argue that it is incidental in *all* cases of noun incorporation (or there would likely be no noun incorporation at all).<sup>29</sup> What we are then left with is the IN representing known information, which is the most concrete criterion, and is the parameter we use in our pilot study.

The items chosen for our pilot study, *bird-watch* and *baby-sit*, were selected because their incorporated nouns have a good chance to represent known information. For all these examples, we checked longer stretches of preceding text. At first sight, quite a number of instances mention birds or children in the previous co-text, for *bird-watching* and *baby-sitting* respectively. However, many of these cannot be analysed as type III, for different reasons.

First, in many examples, the incorporated noun is not equivalent to the previously mentioned entities, and cannot be said to represent ‘known’ information. Instead, a different set of

---

<sup>29</sup> The criterion of representing incidental information is likely why Feist (2013: 168) argues that examples such as (i) instantiate type III, viz. when an indirect participant is incorporated and the resulting predicate is intransitive.

(i) While mantas **filter-feed** in this murky broth of plankton, yellow-striped jacks find protection among the giants. (Feist 2013: 168, from COCA)

To us, however, this does not provide sufficient evidence for type III, for precisely the reason mentioned: all noun incorporation to some extent incorporates incidental information. Since the IN does not represent known information either (we checked the preceding co-text), it is not clear to us why (i) would be specifically different from other type I examples.

referents is implied, as in (41), where the specific babies mentioned first are different from the larger set of babies implied in *baby-sit*, viz. any or all the babies living at Elisabeth House.

- (41) “Give him some kisses,” 17-year-old Cherish McAbee said as her 18-month-old daughter snuggled with the baby of another resident. The girls, who were all homeless, pregnant or young mothers when they were accepted at Elisabeth House in rural Auburn, help each other, *baby-sit* for each other and learn from each other as well. (WB)

Second, some examples also have an earlier mention of specific birds or children, but then the incorporated noun is generic and does not imply any specific referent. In (42), for instance, *three royal spoonbills* are referred to; however, the incorporated *bird* is generic, and the entire N-V compound also refers to the activity of bird-watching generically, and not to the one-off event of watching these three birds, or that bird species, specifically.

- (42) That sense of enjoyment that day was enhanced only a few minutes later when I saw three royal spoonbills perched elegantly, high up in a dead tree. You’re more likely to see these birds wading along the edges of a lake or creek with their aptly named bills searching through the water for food. The unexpected is one of the great joys of *birdwatching*, whether it be the first sight of a new bird or unusual or enchanting behaviour from a more common species. (WB)

Leaving out all such cases, we are left with 3 instances of *bird-watch* and 15 of *baby-sit* (analysed as type I in Figure 2 above) where a referent is first established and then implied by the incorporated noun, which thus seem good candidates for type III incorporation. Example (43a) discusses the migration of cranes, and the bird-watching mentioned clearly involves watching this specific type of birds, and more specifically, a subset of the 15,000 who will fly over the country (even if there is no complete overlap between the reference sets involved). Similarly, in (43b), the incorporated noun *baby* clearly implies the new baby the speaker had, which is mentioned previously.

- (43) a. At the start of this month, sightings were reported of about 400 cranes with wing spans of more than seven feet, mostly using cornfields for a feeding and resting stopover in their migratory journey from wintering grounds in California to breeding and summering sites in Alaska. The crane numbers will typically build to a peak of nearly 15,000 by April 1, which is about half of all the sandhill cranes in the Pacific flyway. Activities at the festival include *bird-watching* field trips, burrowing-owl tours and boat tours at Potholes Reservoir. (WB)
- b. “It was heartbreaking not to be there but we have just had a new baby and couldn’t get anyone to *babysit*.” (WB)

Note that another possible instance of type III incorporation (with a different N-V compound) was discussed in Section 3.1.3, example (15a), repeated here as (44). In this stretch of discourse, *Blaine* is a topical referent, who is then backgrounded and implied by the incorporated noun. As remarked by Mithun, incorporated nouns in type III incorporation are “not necessarily non-specific and indefinite” (1984: 861), which explains why proper names are not excluded from being incorporable.

- (44) So perhaps Londoners are more in the grip of Blaine than he is of them, in spite of all the abuse that he is taking. In that corner of central London, there is a bizarre circus geared around his every need. While I was *Blaine-watching*, there was a surreal argument between Live TV representatives and one of Blaine’s bodyguards. (WB)

While making a good case for type III, these examples still seem different from the ones discussed by Mithun (1984). This is because the incorporated noun cannot simply incorporate into any verb fitting in the context because of lexical restrictions. New coinages like *Blaine-watching*, for instance, seem to always have analogies with existing, well-established compounds like *bird-watching*, and we do not expect *Blaine* to incorporate into many other verbs (as confirmed by a quick search in WB). By contrast, ‘true’ instances of type III (like the examples in Mithun’s work) would show the same incorporated noun combined with different verb stems, as in the made-up story in (45).

- (45) ? The neighbours have a baby girl Sarah. Yesterday I went to their place to baby-feed, baby-wash and baby-sit. Although I was baby-watching all the time, I could not keep her from drawing on the wall. (invented example)

This way of keeping a topic straight over a stretch of discourse is not possible in English (even if 16 different combinations of *baby-V* occur in the corpus). In English, as in other analytic languages, we would use free personal pronouns to refer back to the discourse participant introduced in the first sentence. In many of the (poly)synthetic languages discussed by Mithun (1984: 859), in contrast, a clause does not need any external noun phrases to be grammatical. In fact, such noun phrases are only used to introduce new referents or express focused referents, while topical core arguments are typically indexed on the verb, if overtly expressed at all. In such a system, noun incorporation is a useful device that can – but need not – be used to keep track of a discourse participant across sentences, by “indicating the type of patient, instrument, or location involved in the action or state” (Mithun 1984: 860).

#### 4.4 Interim conclusion

In this section, we discussed whether Mithun’s (1984) typology of noun incorporation can be applied to English. Analysis of our two datasets shows that types I, II and IV are without a doubt available in English. While type I is by far the most common, types II and IV are also well-established, and found with a number of different N-V combinations. In contrast to the examples discussed by Mithun, type IV in English does not lead to a broad system of nominal classification; however, there is some evidence for the availability of miniature, incipient classification systems. Type III, finally, is less clearly established in English. First, its identification was complicated by some analytical difficulties with type III in general, which is only functionally and not structurally distinguished from the other types. Second, in English, unlike in the examples discussed by Mithun, noun incorporation cannot be used to keep a topic straight over a stretch of text, and it seems that not any noun can incorporate into any verb. All type III examples in the corpus are lexically and discursively very specific. As an explanation of these findings, we invoked the analytic nature of the English language, which requires overtly expressed external arguments. The repercussions of this more restricted manifestation of type III for Mithun’s implicational hierarchy of noun incorporation are discussed in Section 5.

### 5. Concluding discussion

This study used two datasets compiled from the *WordBanks Online* corpus to study noun incorporation in English from a usage-based and typological perspective. The first dataset comprised random samples of pre-selected N-V compounds, whereas the second relied on

semi-open searches to capture the full range of possible N-V combinations for specific nouns and verbs. We followed Feist (2013) in using a constructional definition of noun incorporation, viz. as noun-verb compounds resulting in a verb. We thus disregarded their etymology, because this would have a priori limited the scope of our investigation, especially obscuring the role of analogy in the coinage of N-V combinations. First, we showed that noun incorporation in English is, *pace* Mithun (1984), a highly productive process, confirming earlier studies on English by Hall (1956), Rice & Prideaux (1991) and Feist (2013). Overall, noun incorporation is a highly extensible process, with high type frequency (including many low-frequency items alongside a small set of high-frequency items) and a broad variety of semantic relations between the compound parts (in line with cross-linguistic tendencies in Lehmann and Verhoeven (2005)). Our data also revealed cases where the incorporated element is not an uninflected common noun, but a proper noun, inflected noun or nominal complex. The latter two had gone unnoticed in the literature on English noun incorporation so far, whereas they had been observed for other languages. In addition, we found all possible inflectional and derivational verb forms, with many lexemes showing finite forms, including low-frequency items. Incorporated verbs are hence certainly not restricted to occurring as participles or gerunds (as suggested by Mithun [1984: 847] and Rice & Prideaux [1991]), even if these are the most commonly found forms. Of course, language change may have been at work; finite forms may have become more widely used than they were in the 1980s and 1990s (though note that Hall [1956] already mentioned finite usages). Finally, and importantly, our second dataset included many items that are not likely coined through backformation, further strengthening our case for the productivity of noun incorporation in English.

Second, we showed that all four types of noun incorporation proposed by Mithun (1984) in her typological study are found in English, but type III is very restricted and type IV is different from the cases discussed in her paper in that it does not lead to a broad system of nominal classification. We argued that these differences are related to the general typological profile of English, viz. as an analytic language with overtly expressed external arguments, contrary to the polysynthetic languages discussed in Mithun's paper. Despite these differences, we assessed the evidence for type IV strong enough to establish the type confidently for English, but we are less confident for type III. In other words, the implicational hierarchy proposed by Mithun, where languages that have type IV, also have type III, II and I, still holds for English, but should be nuanced to fit the typological profile of more analytic languages. Given this implicational hierarchy, the availability of types other than type I also provides further evidence for the productivity of the phenomenon in English: if unproductive, we would expect noun incorporation to be limited to the most basic type, lexical compounding. Finally, on a methodological note, we discussed how the recognition criteria for the four types can be applied to the wealth of inflectional and derivational forms of compound lexemes observed in English. We hope this discussion will inform future work on noun incorporation in analytic languages.

In addition, our study further extends Mithun's (1984) typology with respect to the nature of the classificatory relation in instances of type IV incorporation. Mithun's discussion features examples in which the classificatory relation resides in inherent, intrinsic characteristics of the referent in the external noun phrase, such as its shape (e.g. spherical shape of incorporated 'eye' for the external noun 'bead'), substance (e.g. incorporated 'granular substance' for the external noun 'parched corn'), or (sub)type specification (e.g. incorporated 'tree' for the external noun 'cashew tree') (Mithun 1984: 865–867). Such examples are indeed found in English as well, for example with *pet-sit* or *house-hunt*, but our data also include examples in which the classificatory relation rather originates in the speaker's or agent participant's attitude. For instance, in *he hero-worshipped Stalin*, Stalin is characterised as a hero by the *he*-agent, but Stalin does not inherently belong to the class of heroes. Put differently, in cases of

non-inherent classificatory relations, speakers might disagree on the validity of the classificatory relation expressed. We believe that this does not impinge on the analysis of such cases as instantiating type IV incorporation; it is just that, to our knowledge, such cases have not been discussed yet. Our study thus raises the question as to whether non-inherent classificatory relations are restricted to analytic languages or can be observed in polysynthetic languages too.

Finally, while this article has focused on the synchronic validity of Mithun's (1984) implicational hierarchy in present-day English, it has left the latter's diachronic implications unaddressed. More concretely, Mithun (1984: 874–877) argues that the development of noun incorporation starts from type I and may be extended over type II to types III and IV respectively. In another study (see Louagie & Van linden (2022); Van linden & Louagie (in prep.)), we investigate this suggested pathway for English, which has good diachronic sources available. We trace the development of a selected set of lexemes, including *baby-sit* and *hand-hold*, through the Corpus of Historical American English. Preliminary results show that the hierarchy largely holds, with type I uses preceding type II and type IV uses. An exception is again type III, which is found only in a restricted way in the corpus, thus confirming the results of the current synchronic study. This study also shows that the backformation story is not always confirmed by the Oxford English Dictionary, with nouns and verbs starting to occur at around the same time. We keep detailed study of the diachronic validity of Mithun's (1984) implicational hierarchy for future work.

## Abbreviations

3 – third person; AN – animate; NCOMP – non-completive; IND – indicative; IPFV – imperfective; NMLZ – nominalizer; PROG – progressive; REC.PST – recent past; TR – transitive; SG – singular

## Data availability statement

The dataset underlying this article can be found on the ULiège Dataverse, the institutional FAIR data repository of the University of Liège: <https://doi.org/10.58119/ULG/AAXPGD>.

## Funding information

Dana Louagie acknowledges the financial support of the Fund for Scientific Research (F.R.S.-FNRS; postdoctoral fellowship).

## Acknowledgements

Noun incorporation in English formed the topic of study in one of An Van linden's MA courses at the University of Liège, and we thank all students who participated in this course for their questions and enthusiasm (2017-2019). While their coursework has not found its way into this study directly, it sparked our curiosity and raised new questions to investigate. We also want to thank Hendrik De Smet for our inspiring discussion on productivity in languages, and Lachlan Mackenzie and two anonymous reviewers for their feedback on an earlier version of this manuscript, which pushed us to add more interaction with existing studies and more descriptive precision in our analyses.

## References

- Aikhenvald, Alexandra Y. 2007. Typological distinctions in word-formation. In Timothy Shopen (ed.), *Language typology and syntactic description, second edition. Volume III: Grammatical categories and the lexicon*, 1–65. Cambridge: CUP.
- Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- Baker, Mark. 1996. *The polysynthesis parameter*. New York: OUP.
- Barðdal, Jóhanna. 2008. Productivity: Evidence from case and argument structure in Icelandic. Amsterdam: John Benjamins.
- Barrie, Michael. 2011. Dynamic Antisymmetry and the syntax of noun incorporation. Dordrecht: Springer.
- Barrie, Michael & Bettina Spreng. 2009. Noun incorporation and the progressive in German. *Lingua* 119. 374–388.
- Barrie, Michael & Audrey Li. 2015. The semantics of (pseudo) incorporation and case. In Olga Borik & Berit Gehrke (eds.), *The syntax and semantics of pseudo-incorporation*, 159–188. Leiden: Brill.
- Booij, Geert. 2008. Pseudo-incorporation in Dutch. *Groninger Arbeiten zur Germanistischen Linguistik* 46. 3–26.
- Borik, Olga & Berit Gehrke. 2015. An Introduction to the Syntax and Semantics of Pseudo-Incorporation. In Olga Borik & Berit Gehrke (eds.), *The syntax and semantics of pseudo-incorporation*, 1–43. Leiden: Brill.
- Bricker, Victoria. 1978. Antipassive constructions in Yucatec Maya. In Nora C. England (ed.), *Papers in Mayan linguistics*, 3–24. Columbia: University of Missouri.
- Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10. 425–455.
- Chafe, Wallace. 1977. Caddo texts. In Douglas R. Parks (ed.), *Caddoan texts*, 27–43. Chicago: University of Chicago Press.
- Collins WordbanksOnline, HarperCollins. <http://www.collinslanguage.com/content-solutions/wordbanks>.
- Dahl, Östen. 2004. The growth and maintenance of linguistic complexity. Amsterdam: John Benjamins.
- Davidse, Kristin & Tine Breban. 2019. A cognitive-functional approach to the order of adjectives in the English noun phrase. *Linguistics* 57(2). 327–371.
- De Smet, Hendrik. 2020. What predicts productivity? Theory meets individuals. *Cognitive Linguistics* 31(2). 251–278.
- De Smet, Hendrik, Lobke Ghesquière & Freek Van de Velde (eds.). 2015. *On multiple source constructions in language change*. Amsterdam: John Benjamins.
- Dixon, R. M. W. 1988. *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- Ebert, Karin H. 2000. Progressive markers in Germanic languages. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe*, 605–653. Amsterdam: John Benjamins.
- Feist, Jim. 2013. Noun incorporation in English. *Functions of Language* 20(2). 159–184.
- Feist, Jim. 2016. *Semantic structure in English*. Amsterdam/Philadelphia: John Benjamins.
- Gerds, Donna B. 1998. Incorporation. In Andrew Spencer & Arnold M. Zwicky (eds.), *The Handbook of Morphology*, 84–100. Oxford and Malden, MA: Blackwell.
- Hall, Robert A., Jr. 1956. How we noun-incorporate in English. *American Speech* 31(2). 83–88.
- Halliday, M. A. K. 1985. *An introduction to Functional Grammar*. London: Arnold.
- Haspelmath, Martin. 2002. *Understanding Morphology*. London: Arnold.

- Haspelmath, Martin. 2024. Compound and incorporation constructions as combinations of unexpandable roots. Unpublished manuscript. Available on <https://ling.auf.net/lingbuzz/008166> [accessed 13 December 2024].
- Lehmann, Christian & Elisabeth Verhoeven. 2005. Noun incorporation and participation. A typological study on participant association with particular reference to Yucatec Maya. In Christian Lehmann (ed.), *Typological studies in participation*, 105–188. Berlin: Akademie.
- Louagie, Dana & An Van linden. 2022. Noun incorporation in English: Diachronic case studies. Paper presented at the Belgian Linguists' Day, Liège.
- Massam, Diane. 2009. Noun incorporation: Essentials and extensions. *Language and Linguistics Compass* 3(4). 1076–1096.
- McGregor, William. 1997. *Semiotic grammar*. Oxford: Clarendon Press.
- McKenzie, Andrew. 2022. Mediating functions and the semantics of noun incorporation. *Natural Language & Linguistic Theory* 40, 505–540.
- Merlan, Francesca. 1976. Noun incorporation and discourse reference in modern Nahuatl. *International Journal of American Linguistics* 42. 177–191.
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60 (4). 847–894.
- Mithun, Marianne. 2000. Incorporation. In Geert Booij, Christian Lehmann, Joachim Mugdan, Wolfgang Kesselheim & Stavros Skopeteas (eds.), *1. Halbband: Ein internationales Handbuch zur Flexion und Wortbildung*, 916–928. Berlin: Mouton.
- Olthof, Marieke. 2020. Formal variation in incorporation: a typological study and a unified approach. *Linguistics* 58(1). 131–205.
- Olthof, Marieke, Eva van Lier, Tjeu Claessen, Swintha Danielsen, Katharina Haude, Nico Lehmann, Maarten Mous, Elisabeth Verhoeven, Eline Visser, Marine Vuillermet & Arok Wolvengrey. 2021. Verb-based restrictions on noun incorporation across languages. *Linguistic Typology* 25(2). 211–256.
- Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York: Basic Books.
- Plag, Ingo. 2003. *Word-formation in English*. Cambridge: CUP.
- Rice, Sally & Gary Prideaux. 1991. Event-packing: The case of object incorporation in English. In *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General session and parasession on the grammar of event structure*, 283–298.
- Rodrigues, Aryon Dall'Igna. Manuscript. *Estrutura do Tupinamba*. Universidade Estadual Campinas, Brazil.
- Rosen, Sara Thomas. 1989. Two types of noun incorporation: A lexical analysis. *Language* 65(2). 294–317.
- Sapir, Edward. 1911. The problem of noun incorporation in American languages. *American Anthropologist* 13(2). 250–282.
- Smirnova, Elizaveta & Svetlana Shustova. 2017. Noun-incorporation in English as a valency-changing device. *Lingua* 194. 15–25.
- Stvan, Laurel Smith. 2009. Semantic incorporation as an account for some bare singular count noun uses in English. *Lingua* 119. 314–333.
- Van linden, An. 2023. Harakmbut. In Patience Epps and Lev Michael (eds.), *Amazonian Languages, An International Handbook. Language Isolates*, Volume 1: Aikanã to Kandozi-Shapra, 441–481. Berlin: Mouton.
- Van linden, An & Dana Louagie. In preparation. Noun incorporation in English: A diachronic perspective.
- Weggelaar, C. 1986. Noun incorporation in Dutch. *International Journal of American Linguistics* 52(3). 301–305.

**Author details**

Department of Modern Languages

place Cockerill 3-5

4000 Liège

Belgium

[dana.louagie@uliege.be](mailto:dana.louagie@uliege.be); [an.vanlinden@uliege.be](mailto:an.vanlinden@uliege.be)

## Appendix

**Table i.** N-V combinations in Dataset 2

N-	-V	tokens	finite tokens	semantic role	type Mithun (1984)
N-randomV					
baby	boost	1	0	patient	I
baby	butcher	1	0	patient	I
baby	change	2	0	patient	I
baby	dangle	1	0	patient	I
baby	eat	1	0	patient	I
baby	feed	2	0	patient	I
baby	handle	1	0	similative	IV
baby	kill	2	0	patient	I
baby	kiss	1	0	patient	I
baby	make	3	0	patient	I
baby	mind	1	0	patient	I
baby	name	3	0	patient	I
baby	proof	6	1	beneficiary	I
baby	tend	1	0	patient	I
baby	watch	1	0	patient	IV
bird	attract	1	0	patient	I
bird	band	1	0	patient	I
bird	bite	1	0	patient	I
bird	collect	1	0	patient	I
bird	eat	6	0	patient	I
bird	love	1	0	patient	I
bird	ring	1	0	patient	I
brain	addle	2	0	patient	I
brain	blunt	1	0	patient	I
brain	build	1	0	patient	I
brain	bust	1	0	patient	I
brain	clog	1	0	patient	I

brain	crumble	1	0	patient	I
brain	crunch	1	0	patient	I
brain	crush	3	0	patient	I
brain	damage	92	1	patient	I; II
brain	destroy	2	0	patient	I
brain	dizzy	1	0	patient	I
brain	drain	2	0	patient	I; II
brain	dull	1	0	patient	I
brain	eat	1	0	patient	I
brain	fix	1	1	patient	I
brain	fry	1	0	patient	I
brain	image	7	0	patient	I
brain	injure	1	0	patient	II
brain	liquefy	1	0	patient	I
brain	lock	1	1	patient	I
brain	mangle	1	0	patient	I
brain	map	3	0	location; patient	I
brain	melt	1	0	patient	I
brain	monitor	1	0	patient	I
brain	munch	1	0	patient	I
brain	nourish	2	0	patient	I
brain	numb	4	0	patient	I
brain	optimise	1	0	patient	I
brain	paralyse	1	0	patient	I
brain	rattle	2	1	patient	I; II
brain	save	1	0	patient	I
brain	scan	5	0	patient	I
brain	scramble	1	0	patient	I
brain	signal	1	0	location	I
brain	swell	1	0	patient - S	I
brain	tan	1	0	instrument	I
brain	tease	2	0	patient	I

brain	wash	529	56	patient	I; II; IV
brain	waste	14	0	patient	I
foot	bind	2	0	patient	I; II
foot	clog	1	0	patient	I
foot	clop	1	0	manner	I
foot	crunch	1	1	manner	I
foot	drag	6	0	patient	I
foot	hold	1	0	patient	I
foot	lash	1	0	patient	II
foot	lift	1	0	patient	I
foot	operate	1	0	manner	I
foot	rush	1	1	manner	I
foot	slog	6	1	manner	I
foot	stamp	4	0	manner	I
foot	stomp	7	1	manner; instrument	I
foot	tap	5	0	manner	I
foot	wash	1	0	patient	I
hand	address	3	1	manner	I
hand	apply	1	0	manner	I
hand	assemble	2	1	manner	I
hand	bail	1	1	manner	I
hand	batik	1	0	manner	I
hand	bead	3	1	manner	I
hand	beat	1	0	manner	I
hand	bend	1	1	manner	I
hand	blend	1	0	manner	I
hand	blow	1	0	manner	I
hand	brake	1	1	manner	I
hand	build	9	0	manner	I
hand	card	1	0	manner	I
hand	carry	4	0	manner	I
hand	carve	11	2	manner	I

hand	cast	2	1	manner	I
hand	chase	1	0	manner	I
hand	check	2	0	manner	I
hand	choose	2	0	manner	I
hand	chrome	1	0	manner	I
hand	clap	8	2	manner	I; II
hand	clasp	2	0	patient	I
hand	clean	1	0	manner	I
hand	coat	1	0	manner	I
hand	code	1	0	manner	I
hand	collect	1	0	manner	I
hand	colour/color	5	0	manner	I
hand	compress	1	0	patient	I
hand	count	10	0	manner	I
hand	cover	3	1	manner; instrument	I
hand	craft	224	3	manner	I
hand	crank	1	0	manner	I
hand	create	1	0	manner	I
hand	cripple	1	0	patient	I
hand	cuff	1311	216	patient	I; II; IV
hand	cultivate	1	0	manner	I
hand	cut	2	0	manner	I
hand	date	2	1	manner	I
hand	decorate	5	0	manner	I
hand	deliver	20	1	manner	I
hand	develop	1	0	manner	I
hand	dip	4	0	manner	I
hand	distil	1	0	manner	I
hand	dovetail	1	0	manner	I
hand	draft	1	0	manner	I
hand	drop	1	0	manner	I
hand	dry	1	1	manner	I

hand	embellish	1	0	manner	I
hand	emboss	1	0	manner	I
hand	embroider	1	0	manner	I
hand	enamel	1	0	manner	I
hand	engrave	2	0	manner	I
hand	express	1	1	manner	I
hand	fashion	1	0	manner	I
hand	feed	38	12	manner	I
hand	fertilize	1	0	manner	I
hand	file	2	0	manner	I
hand	finish	13	0	manner	I
hand	flap	1	0	patient	I
hand	fly	1	0	manner	I
hand	forge	6	3	manner	I
hand	form	3	0	manner	I
hand	gather	1	0	manner	I
hand	gild	2	0	manner	I
hand	graft	1	0	manner	I
hand	graph	1	1	manner	I
hand	hammer	1	0	manner	I
hand	harvest	3	0	manner	I
hand	hew	19	0	manner	I
hand	infuse	1	0	manner	I
hand	ink	1	0	manner	I
hand	inscribe	1	0	manner	I
hand	inspect	2	1	manner	I
hand	knit	51	1	manner	I
hand	knot	10	0	manner	I
hand	label	1	0	manner	I
hand	lace	1	0	manner	I
hand	landscape	1	0	manner	I
hand	layer	1	0	manner	I

hand	letter	1	0	manner	I
hand	link	1	0	location	I
hand	loom	1	0	manner	I
hand	make	6	1	manner	I
hand	massage	1	0	manner	I
hand	mend	1	1	manner	I
hand	milk	2	0	manner	I
hand	mix	2	2	manner	I
hand	model	1	0	manner	I
hand	nail	1	1	manner	I
hand	number	5	0	manner	I
hand	pack	3	0	manner	I
hand	paint	92	5	manner	I
hand	perform	1	0	manner	I
hand	pick	416	57	manner	I
hand	pinch	1	0	manner	I
hand	place	3	1	manner	I
hand	plane	1	1	manner	I
hand	plant	1	0	manner	I
hand	plot	1	0	manner	I
hand	plunge	1	0	manner	I
hand	polish	2	0	manner	I
hand	pollinate	4	2	manner	I
hand	pour	1	0	manner	I
hand	prepare	1	0	manner	I
hand	press	2	1	manner	I
hand	print	13	3	manner	I
hand	pump	5	1	manner; patient	I
hand	quilt	2	0	manner	I
hand	raise	7	0	manner	I
hand	rear	20	4	manner	I
hand	rip	1	0	patient	I

hand	roast	1	0	manner	I
hand	roll	16	0	manner	I
hand	rotate	1	0	manner	I
hand	rub	28	1	manner; patient	I
hand	salt	1	0	manner	I
hand	sanitize	1	0	patient	I
hand	saw	1	1	manner	I
hand	scrape	1	0	manner	I
hand	scratch	1	1	manner	I
hand	scrawl	1	0	manner	I
hand	screen	3	1	manner	I
hand	scribble	1	0	manner	I
hand	scribe	1	0	manner	I
hand	scrub	2	0	manner	I
hand	sculpt	2	0	manner	I
hand	search	7	1	manner	I
hand	select	10	4	manner	I
hand	sell	1	1	manner	I
hand	sew	56	9	manner	I
hand	shake	1	1	patient	I
hand	sift	1	0	manner	I
hand	sign	4	0	manner	I
hand	signal	2	1	instrument	I
hand	sketch	1	0	manner	I
hand	slice	1	1	manner	I
hand	slip	1	1	location	I
hand	sort	4	0	manner	I
hand	spray	5	1	manner	I
hand	spring	1	0	manner	I
hand	stain	1	0	manner	I
hand	stamp	2	0	manner	I
hand	steer	3	0	manner	I

hand	stencil	1	0	manner	I
hand	stir	1	0	instrument	I
hand	stitch	18	3	manner	I
hand	stress	1	0	manner	I
hand	strip	2	0	manner	I
hand	stroke	1	0	patient	I
hand	structure	1	0	manner	I
hand	tailor	2	0	manner	I
hand	tie	3	1	manner; patient	I; II
hand	time	3	0	manner	I
hand	tint	2	0	manner	I
hand	tool	1	0	manner	I
hand	tow	1	0	manner	I
hand	trace	1	0	manner	I
hand	treat	1	0	manner	I
hand	tremble	2	0	patient - S	I
hand	trip	2	0	manner	I
hand	turn	2	0	manner	I
hand	vet	1	0	manner	I
hand	walk	2	2	manner	I
hand	warm	1	0	manner	I
hand	wash	29	5	manner; patient	I
hand	water	2	0	manner	I
hand	wave	4	1	manner	I
hand	wax	1	0	manner	I
hand	weave	5	0	manner	I
hand	weed	3	0	manner	I
hand	weld	1	1	manner	I
hand	work	1	1	manner	I
hand	wrap	2	0	manner	I
hand	wrestle	1	0	manner	I
hand	write	43	2	manner	I

head	bang	38	1	patient	I; II
head	bash	1	1	patient	II
head	batter	1	0	patient	I
head	bob	1	0	patient	I
head	bop	1	0	manner	I
head	bury	1	0	patient	I
head	butt	271	122	manner	I
head	catch	1	0	instrument	I
head	clean	1	0	patient	I
head	clear	3	0	patient	I
head	cushion	1	0	patient	I
head	dizzy	1	0	patient	I
head	flick	1	1	instrument	I
head	hammer	1	0	patient	I
head	hug	1	0	patient	I
head	lay	1	1	patient	I
head	mount	2	0	location	I
head	nod	3	0	manner	I
head	protect	8	0	patient	I
head	reel	1	0	patient	I
head	scratch	10	0	patient	I
head	seal	1	0	patient	I
head	shake	9	2	patient	I
head	shave	1	0	patient	I
head	slam	1	0	patient	II
head	smash	2	1	patient	I; II
head	snap	1	0	patient	I
head	spin	10	0	patient - S	I
head	thump	1	0	patient	I
head	turn	15	0	patient	I
head	wiggle	1	0	patient	I
head	wobble	1	0	patient	I

randomN-V					
actor	hunt	1	0	patient	I
apartment	hunt	2	1	patient	I
autograph	hunt	3	0	patient	I
bargain	hunt	1	1	patient	I
Beast	hunt	1	0	patient	I
big-game	hunt	1	0	patient	I
boar	hunt	1	0	patient	I
bow	hunt	1	0	instrument	I
buffalo	hunt	1	0	patient	I
carrot	hunt	1	0	patient	I
cash	hunt	1	0	patient	I
cave	hunt	1	0	patient	I
coon	hunt	1	1	patient	II
croc	hunt	1	0	patient	I
cub	hunt	1	0	patient	I
curtain	hunt	1	0	patient	I
deer	hunt	1	0	patient	I
drag	hunt	8	0	patient	I
elephant	hunt	1	0	patient	I
first world war bomb	hunt	1	0	patient	I
flat	hunt	11	0	patient	I
fortune	hunt	1	0	patient	I
fossil	hunt	2	0	patient	I
fox	hunt	10	0	patient	I
ghost	hunt	4	0	patient	I
girl	hunt	1	0	patient	I
glory	hunt	3	0	patient	I
goose	hunt	1	0	patient	I
human	hunt	1	0	patient	I
husband	hunt	2	0	patient	I

insect	hunt	1	0	patient	I
job	hunt	3	0	patient	I
man	hunt	1	0	patient	I
migrant	hunt	1	0	patient	I
moose	hunt	1	0	patient	I
Nazi	hunt	1	0	patient	I
partner	hunt	1	0	patient	I
planet	hunt	3	0	patient	I
plant	hunt	2	0	patient	I
poacher	hunt	1	0	patient	I
promotion	hunt	2	0	patient	I
property	hunt	1	0	patient	I
rabbit	hunt	1	0	patient	I
room	hunt	1	0	patient	I
scalp	hunt	1	0	patient	I
slave	hunt	1	0	patient	I
souvenir	hunt	3	0	patient	I
sponsor	hunt	1	0	patient	I
spy	hunt	1	0	patient	I
stag	hunt	2	0	patient	I
submarine	hunt	1	0	patient	I
tank	hunt	1	0	patient	I
treasure	hunt	2	0	patient	I
truffle	hunt	2	0	patient	I
whale	hunt	1	0	patient	I
witch	hunt	3	0	similative	IV
woman	hunt	3	0	patient	I
bench	sit	1	0	location	I
cat	sit	2	0	comitative	I
cat or dog	sit	1	0	comitative	I
dog	sit	6	1	comitative	I
egg	sit	1	0	location-comitative	I

fence	sit	25	1	location	I
flat	sit	1	0	location-comitative	IV
ground	sit	1	0	location	I
house	sit	48	3	location-comitative	I; IV
lap	sit	1	0	location	I
pet	sit	5	0	comitative	I; IV
phone	sit	1	0	comitative	I
pillar	sit	1	0	location	I
pole	sit	6	0	location	I
sauna	sit	1	0	location	I
sidewalk	sit	1	0	location	I
site	sit	1	1	location-comitative	IV
stair	sit	1	1	location	I
stool	sit	1	0	location	I
stoop	sit	1	0	location	I
tree	sit	1	0	location-comitative	I
villa	sit	1	0	location-comitative	I
window	sit	1	0	location	I
ant	watch	1	0	patient	I
badger	watch	1	0	patient	I
Blaine	watch	1	0	patient	I
breath	watch	1	0	patient	I
budget	watch	1	0	patient	I
carousel	watch	1	0	patient	I
clock	watch	2	2	patient	I
dolphin	watch	2	0	patient	I
film	watch	1	0	patient	I
fire	watch	2	2	patient	I
fundamentals	watch	1	0	patient	I
game	watch	1	0	patient	I
gorilla	watch	1	0	patient	I
hawk	watch	1	0	patient	I

kid	watch	1	0	patient	I
marriage	watch	1	0	patient	I
movie	watch	3	0	patient	I
news	watch	1	1	patient	I
people	watch	25	1	patient	I; IV
politician	watch	1	0	patient	I
scoreboard	watch	2	0	location; patient	I
sky	watch	1	0	patient	I
squirrel	watch	1	0	patient	I
star	watch	4	1	patient	I
TV	watch	1	0	patient	I
waste	watch	1	0	patient	I
weight	watch	4	0	patient	I
whale	watch	18	0	patient	I
Total		4196	584		