

# N+N Structures in Present-Day English Word Formation<sup>1</sup>

## Iria Pastor Gómez

*University of Santiago de Compostela*

### Abstract

N+N structures are sequences of two nouns, such as *drug addiction* or *heart attack*; often, structures of this type appear to be formations to which speakers and writers resort on the spur of the moment in order to fill a semantic gap, and which will never be used again. However, reasons such as their relevance or their easiness to be understood help them to remain through time as structures which enter the everyday vocabulary of a given language. The present paper is an approach to the consideration that N+N structures are a productive word formation type in Present-Day English. For this reason, I will try to establish which the motivations for their use are and I will make reference to the process of lexicalisation they may undergo. Likewise, I will make use of a corpus of English written texts to illustrate the lexical richness that these formations may provide to the vocabulary of English. The data taken from the corpora show that there is an active progress in the development of new compositional nominal groups.

## 1. Introduction

This paper is a preliminary approach to the analysis of N+N structures as a productive word formation type in Present-Day English. N+N structures are sequences of two nouns, such as *drug addiction* or

---

<sup>1</sup> This paper has been done with the support of the research project *Variation, Linguistic Change and Grammaticalization*, funded by the Spanish Ministry of Education and Science (grant no. HUM2007-60706 (CONSOLIDER)), the Autonomous Government of Galicia (grants no. PGIDIT05PXIC20401PN, 2006/XA-133) and the European Regional Development Fund (ERDF). Support from the Spanish Ministry of Education and Science through the research fellowship AP2004-4917 is also gratefully acknowledged.

## PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE

*heart attack*; often, structures of this type appear to be formations to which speakers and writers resort on the spur of the moment in order to fill a semantic gap, and which will never be used again. Yet a number of factors may cause some of these formations to become entrenched and thus become part of the everyday vocabulary of English as genuine compounds.

In what follows I will try to determine which are the motivations behind the use of N+N structures, and I will also examine some of the idiosyncratic features they acquire in the process of lexicalisation. My analysis makes use of a corpus of English written texts in order to show the important contribution of these formations to the vocabulary of English.

### 2. Motivations behind the use of N+N structures

It is well-known that speakers and writers may use vocabulary and syntax in such a way that they can give rise to a high number of different combinations which can convey whatever meaning they intend to. However, clauses may be replaced, most of the time, by shorter constructions such as N+N structures:

- (1a) a paper made of linen fibres
- (1b) linen paper

Thus, the clause in (1a) can be replaced by a shorter sequence of nouns in (1b).

The question which follows is why there is a need to express concepts by means of a sequence of nouns. First of all, it may be adduced that N+N structures are a compacted device which can be used for linguistic economy motivations, given that N+N structures provide with a reduced expression of what can be said with a longer sentence. Press language may serve as an example: journalists need to deliver as much information as possible with the lowest number of words since space limitations are one of the main constraints in papers and magazines. The role of N+N structures is thus justified, since they represent expressive, coherent and non repetitive forms based on a previous knowledge, which provide the text with a unitary character.

This can also be related to what Haiman (1985: 133) understands by “economic motivation”, by which “the length of linguistic structures will be inversely proportional to their frequency”, i.e., we use shorter structures when the ideas gathered within them are frequently employed by speakers. Likewise, Baayen & Renouf (1996: 141) refer to the concept of *availability*, and they point out that “the more familiar a concept is, the more available it is.” This availability is related to the interlocutors’ previous information about the concept which has been delivered in prior conversation or it is part of the common knowledge a community of speakers has about that concept. Likewise, Varantola (1993: 72) states that “naturally, the lack of verbs and prepositions in premodified noun phrases leads to a loss of explicit textual information and requires more shared knowledge between the writer and the readers than an expression containing a verb and/or postmodifying constructions.” Thus, there is no need to repeat the whole of a sentence in order to refer to an idea which is already known by the interlocutors. The speaker/writer will then use the smallest number of words in order to make such a reference. In relation to this, Baayen & Renouf (1996: 142) also employ the term *predictability*, which is defined as “[...] the degree to which the listener can predict the presence and role of a particular concept.” Thus, we are aware of what is already known by both participants in the linguistic interchange and of what has to be mentioned and omitted to know what the speaker/writer refers to. Furthermore, N+N structures may be used depending on the kind of reaction we want to cause in our audience. Along this line, Varantola points out:

Premodification is a functional means of creating compact packages of information. It is space-saving and non-redundant; it captures a static, crystallised image of an extralinguistic object, event or process shared by a set of readers. It makes no difference whether this set consists of the general public in a particular society or a group of specialists. (1993: 75)

In contrast, longer sentences and postmodifiers give rise to more dynamic expressions which lose the character of universality and describe, instead of classifying, concepts of reality. Compounds, as

morphological items, are created and used because they have a long life, or at least they are supposed to live a long life and thus to be used many times for many communicative purposes. However, relative clauses, for example, are transient, i.e., they are used for a concrete moment and purpose and we do not expect them to survive in the memory for a new usage. Scientific language may serve as an illustration since N+N structures provide one with a universalistic vision of the concepts they convey, which causes these concepts to become widely accepted by the scientific community.

### **3. Lexicalisation of N+N structures and word formation**

As has been said above, one of the main reasons for the use of N+N structures has to do with economic motivations (especially in journalistic language), given that N+N structures allow the writer or speaker to communicate rapidly using the lowest number of words. N+N structures are then used momentarily and this fact does not entail that they must appear again in similar contexts. Along this line, we may refer to Portero's definition of *nonce formation* as “[...] a new complex word coined by a speaker/writer on the spur of the moment to cover some immediate need” (2004: 71). Bauer (2001: 36) points out that this nonce formation becomes into a word if it is item-familiar “to a large enough subset of the speech community.” As a result, this word becomes institutionalised. This process of institutionalisation is defined as “[...] the stage in the history of a lexeme when it becomes to be accepted by other speakers as a known lexical item” (Portero, 2004: 71). Furthermore, there are various reasons why some words become institutionalised and some others do not: the status of the first user, the clarity and prestige of that complex word, the convenience of the complex word to achieve a concrete effect or to save space and the frequency of the designated object.

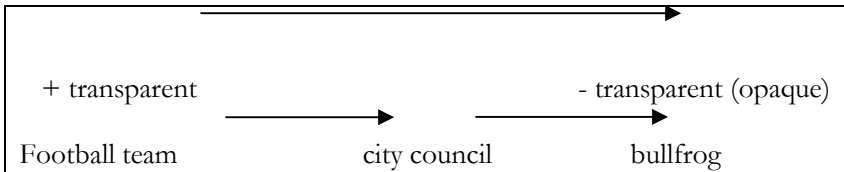
Because of these reasons, some N+N structures tend to become more institutionalised and frequent in time, and, consequently, they may be more specialised in their meaning. Their semantic content is not compositional anymore. As a result, they become more opaque and obscure. This is the process whereby they become lexicalised compound nouns.

Lexicalisation is thus defined by Portero (2004: 72) as “[...] the stage in the history of a vocabulary-item when, because of some change in the language system, the lexeme has, or takes on, a form or a meaning which it could not have had if it had arisen by the application of productive rules.” Lexicalised nominal sequences cannot always be so easily explained using the terms of the lexical items they contain. This is the case of:

(2) home rule

which is not “the rule of a concrete home”, but the “Government of a country, colony, province, etc. by its own citizens” (*OED*).

The existence of this process implies a progression and change through time of a category from being transparent to becoming obscure from the semantic point of view. The difference lies in the fact that there is a higher level of *opacity* which comes about when there is a need for it to be listed in a dictionary in order to understand its meaning. This helps to regard this process as a *continuum* by which N+N structures are in a progress towards their specialisation:



#### 4. Morphological productivity as shown in the corpus

In order to show the lexical richness with which N+N structures contribute within a process of word formation, I have made use of the *Brown, Frown, Lob and Flob Corpora of English Written Texts*, which are a series of compiled passages at the University of Freiburg in the 1990's (see Mair, 2002)<sup>2</sup>. They allow researchers to different investigations according to diachronic and synchronic perspectives. Data from the corpora will show that the number of nominal

<sup>2</sup> Available in Hofland *et al.* (1999).

**PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE**

compounds in English is large enough to state that this is a productive process of word formation. A total of 200,000 words have been checked and consequently many N+N structures have been found (almost 5,000). However, owing to space limitations, this paper just gathers a small collection of the N+N structures found. Tables 1 and 2 collect a compilation of N+N structures which are classified according to the noun from which they derive. They try to display the great variety of compositions that only one noun can give rise to. Thus, in table 1, N+N structures are classified according to the noun which takes the position of the left-hand member. Nominal sequences are classified in table 2 according to the noun that is the rightmost element.

Table 1. Classification of N+N structures according to the left-hand element

AIR	aircraft	Airport	PRESS	press agency	Pressmen
	air crash	air quality		Press Association	press play
	Airfield	air raid		press conference	press room
	air game	Airway		press corps	press secretary
DRUG	drug addiction	drug companies	RADIO	radio activity	radio news
	drug addicts	drug profits		radio announcer	radio noise
	Drug Administration	Drugstores		radio brightness	radio-observations
	drug dealers	drug-traffickers		radio hosts	radio system
GAS	gas constant	gas phase	SEA	sea frontier	Seaports
	gas cooker	gas pressure		sea gastropods	sea-shells
	gas flow	gas samples		sea level	seaside
	gas-glass	gas stations		sea passengers	sea suspension

HEART	heart attack	heart sharers	TAX	Tax Appeals	tax increases
	heartbeat	heartstrings		tax bill	tax payers
	heart break	heart surgery		tax cut	tax policy
	heart disease	heart walls		tax deductibility	tax reform

Table 1 shows a great number of nouns from which many N+N structures derive. The case of *air*, *radio*, or *sea* is especially notable, with a medium rate of 25 N+N structures derived from each of them. Those nouns seem especially productive and give rise to a disparity of meanings which differ much from one another at times. Thus, in the case of *air*, the meaning of *air* in *air quality* (the quality of the air we all breathe) diverges from the meaning of *air* as a locative in *air crash* (a crash originated in the air). Likewise, the majority of the most frequently used nouns are quite generic, that is, their meaning is so general that the addition of another noun to the left-hand element can give rise to a wide variety of new concepts. This is the case of *tax*. In table 1 we note a wide range of new concepts derived from that noun, like *tax increase* (one of the most frequent N+N structures in the corpora), *taxpayers*, etc. However, *tax* behaves in a different way than *air* from a semantic perspective, given that *tax* restricts its meanings to the semantic field of economics while *air*, as has been noted, or *radio* (*radio-emission* vs. *radio news*) exhibit a wider range of meanings. Likewise, there are instances in the corpora of N+N structures with some of these nouns in the right-hand position. This is the case of *fuel-air*, *night air*, or *pressure gas*.

Table 2. Classification of N+N structures according to the right-hand element

AGENCY	education agency	Press Agency	POWER	brain power	missile power
	news agency	Protection agency		emergency powers	people power

PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE

	police agency	transportation agency		fire power	will power
FORCE	Action Force	hurricane forces	SOURCE	energy source	phosphate sources
	air force	market forces		forest sources	power source
	defence force	Volunteer Force		light source	Reference sources
PAPER	ballot papers	newspaper	STATION	battle station	gas stations
	campus paper	toilet paper		bus station	police station
	filter paper	Tory papers		fire stations	TV stations
RATE	burial rates	interest rates	TAX	export tax	poll tax
	cable rate	mortality rate		income taxes	property taxes
	dissipation rate	radiation rates		luxury tax	state tax

Table 2 gathers a compilation of those N+N structures which share the same noun as the right-hand constituent. The number of groups found is not as large as in the case of table 1, nor the number of different N+N structures found for each right-hand constituent. Even so, there are nouns like *rate*, *source* or *station* which show an average number of 10 nominal structures each. Most of these nouns are also quite generic and in many cases the left-hand constituent functions as a classifier of these nouns, that is, the addition of the left-hand constituent gives rise to a new class of N2. *Coffee-shop/pizza shop*, *plane crash/car crash*, or *ice sheet/oxygen sheet* are illustrations of this kind. Also, some nouns in right-hand position may change their meaning depending on the noun they are adhered to, e.g., *filter paper* vs. *newspaper*. Likewise, there are a few right-hand position nouns which can also appear in left-hand position like *rate constants* or *paperwork*. Besides, the case of *trade union* is especially notable, given that the combination of



these two nouns gives rise to a wider range of new composite nouns. The different use of capital letters and plural marks plays a fundamental role for the achievement of new forms like *Trades Union/Trade Union*, *trades union/trade union*, etc.

## 5. Conclusions

As has been already said, nominal structures are used in Present-Day English owing to reasons such as their compactness, which have the capacity indeed to provide a static and even universal character to the concepts they convey. Furthermore, reasons like their frequency of use make these sequences of nouns to be institutionalised. They consequently may acquire a more specialised meaning and so become lexical units.

As a result, N+N structures are a word formation type and contribute to enrich the lexical diversity of a given speech community. The data taken from the corpora illustrate that there is an active progress in the development of new compositional nominal groups. The corpora that have been used in the present paper allow working with different variables such as speech community or gender, which could be the purpose for further research on N+N structures.

## References

- Baayen, R. H. and A. Renouf. 1996. "Chronicling The Times: Productive Lexical Innovations in an English Newspaper". *Language* 72. 69–96.
- Bauer, L. 2001. *Morphological Productivity*. Cambridge: CUP.
- Haiman, J. 1985. *Iconicity in Syntax: Proceedings of a Symposium on Iconicity in Syntax, Stanford, June 24–26, 1983*. Amsterdam: John Benjamins.
- Hofland, K., A. Lindebjerg and J. Thunestvedt. Eds. 1999. *ICAME COLLECTION of English Language Corpora*. University of Bergen: The HIT Centre.
- OED= Simpson, J.A. and E.S.C. Weiner. Eds. 1989. *The Oxford English Dictionary on CD-ROM*. Oxford: OUP.
- Portero, C. 2004. *A Course in English Morphology*. Córdoba: Servicio de Publicaciones.
- Varantola, K. 1993. "Modification of nouns by nouns. Bad by definition?" *Anglistik & Englischunterricht. The Noun Phrase in English. Its Structure and*

**PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE**

*Variability.* H. J. Diller, S. Kohl, J. Kornelius and E.G. Stratmann. *Eds.*  
Heidelberg: Winter.