Syntactic ambiguity of (complex) nominal groups in technical English

MIRJANA BORUCINSKY & JANA KEGALJ
University of Rijeka (Croatia)

ABSTRACT
Complex nominal groups are common in technical English (i.e., English for Specific Purposes, ESP) since they allow lexical items to be tightly packed into a clause which consequently leads to increased lexical density and syntactic ambiguity. In this paper, we analyse (complex) nominal groups in technical English. We propose that, in addition to context and extralinguistic knowledge (i.e. shared technical background that the ESP teacher does not necessarily possess), the structure of the nominal group—or, more precisely, the position of modifiers within the group also plays a role in resolving of syntactic ambiguity and disambiguation of meaning. Thus, modifiers standing farthest from the head have the least specifying potential and are followed by other modifiers that restrict the meaning of the entire nominal group. In this way, the participle *reciprocating* in *steam reciprocating engine* (vs.*reciprocating steam engine*) is more specific in meaning and is thus positioned closer to the head of the nominal group. Our results indicate the type of modification (i.e. linear or non-linear) lends support to the disambiguation of complex nominal groups. The paper's main contribution is in the field of ESP teacher education in the way that it helps ESP teachers who are not specialists in the field of (marine) engineering to process understand and successfully teach complex nominal groups.

1. INTRODUCTION

The (complex) nominal group1 is commonly found in scientific discourse (Halliday, 1993). This is a distinct feature of technical English (which is also referred to as English for Specific Purposes, ESP)2 because it allows the 'packing' of lexical items as a syntactic unit into a clause. However, this results in higher lexical density and it gives rise to syntactic ambiguity. Moreover, it is important to distinguish the term *syntactic ambiguity* from the term *syntactic polysemy*, the latter being "the immediate realisation of more than one categorial meaning within the head element of a language structure" (Kozerenko, 2004: 2). Hence, polysemous structures "display variable manifestation of their categorical features

*****Address for correspondence**: Mirjana Borucinsky & Jana Kegalj, Studenska ulica 2, Faculty of Maritime Studies, University of Rijeka, Croatia; e-mail: mborucin@pfri.hr, kegalj@pfri.hr

depending on the functional role in the sentence" (Kozerenko, 2004: 2). Meanwhile, an ambiguous syntactic structure "presupposes alternative ways of interpretation" (Kozerenko, 2004: 2) and, in order to understand and correctly interpret such ambiguous structures (e.g. *the new control system updated continuously displayed robust performance*), a broader context is deemed necessary. In this paper, we use the term syntactic ambiguity since we deal with nominal groups of complex structures that are difficult to process and which can easily be misunderstood or interpreted alternatively.

Nominal groups constitute "a substantial majority of all technical vocabulary" (Justeson & Katz, 1995: 9) and are distinguished from other types of nominal groups. Biber, Stig, Leech, Conrad and Finegan (1999: 578) have shown that "in all registers, noun phrases with premodifiers are somewhat more common than those with postmodifiers." In addition, the ability of the English language to form complex nominal groups is "an efficient means by which dense information content can be packed into as few words as possible" (cf. Zhoulin, 2016: 76), which allows more information to be communicated at once. Furthermore, Halliday (1993) holds that written language is more lexically dense than spoken language and "in much scientific writing, almost all the lexical items in any clause occur inside just one or two nominal groups" (Halliday, 1993: 84). Hence, nominal groups create most difficulty in processing because "they consist of strings of lexical words without any grammatical words in between" (ibid.) (e.g. *piston outlet lubrication oil temperature* 'the temperature of the lubrication oil that is used to cool the piston and which is measured at the outlet'). Although the extralinguistic context helps us to understand complex nominal groups (cf. Kereković, 2016), very specific technical knowledge is required to do so and non-specialists of the field (e.g. ESP teachers) generally do not possess this knowledge. Hence, this paper aims to find an efficient method of processing and analysing complex nominal groups in technical English, specifically those that consist of three or more elements and which exhibit one of the following structures: N1 + N2 + N3 + NN; Adj+ N1 + N2 + NN.

This paper is laid out as follows: Section 2 outlines the theoretical framework that is adopted (i.e. *Systemic Functional Grammar*) and the key concepts pertaining to the analysis. In Section 3, the methodology is explained. Section 4 presents the analysis and discussion. Finally, Section 5 draws the conclusions and outlines the implications for further research.

2. THEORETICAL BACKGROUND

Systemic Functional Grammar (SFG, Fawcett, 2010, 2017; Halliday, 1985, 1994; Halliday & Matthiessen, 2004) is one of the contemporary functional language models which allows us to analyse and interpret complex nominal groups in a simple and efficient way. Within this theoretical framework, language is seen as part of social semiotics (Halliday, 1978) where

the corpus seems to be an appropriate method for analysing and understanding how language is applied or used in a social context. Consequently, a corpus-based approach was chosen for the study.

SFG proposes a set of terms that show how the clause can be broken down into functional constituents; that is, the participant, process, and circumstance. The participant element is realised by a nominal group; that is, the "THING" (Halliday & Matthiessen, 2004: 67; Martin, Matthiessen & Painter, 2010). Similar to a clause, the nominal group in English is a realisation of three distinct functional components that express three largely independent sets of semantic choice: the ideational, the interpersonal and the textual (Halliday & Matthiessen, 2004: 309). In a clause, the three structures combine into one interpretation, whereas beneath the clause (i.e. in a nominal group) the three structures are incomplete in themselves and need to be interpreted separately "as partial contributions to a single structural line" (Halliday & Matthiessen, 2004: 309). The most significant structure for understanding and processing complex nominal groups is the ideational structure, which is split into two metafunctions: the experiential and the logical (Halliday & Matthiessen, 2004: 309). The experiential structure of the nominal group is made up of the following functions: *Deictic, Numerative, Epithet, Classifier, Thing and Qualifier,* whereby the class of determiner most typically realises the function of the deictic, and the class of numeral realises the numerative function. Adjectives can function as either epithets or classifiers, whereas nouns most commonly function as either the *Thing* or classifiers. Prepositional phrases, other nominal groups, and relative clauses function as qualifiers within a nominal group. The most relevant part of Halliday's theory for understanding how elements are ordered within a nominal group is illustrated by the following quote: "the more permanent the attribute of a Thing, the less likely it is to identify it in a particular context" (Halliday & Matthiessen, 2004: 333). The most permanent item is the Thing (the head itself) and the least permanent is the determinative,3 which will be placed farthest from the head. This pattern of progression explains why the elements of the nominal group are ordered in a specific way. Our main focus here lies on epithets and classifiers and how their position determines the meaning of the nominal group and helps to disambiguate complex nominal groups. Epithets serve to describe or provide additional information about the head without affecting the reference of the head, while the classifier combines with the head to form a referring unit (Warren, 2003: 239).4 Given that the function of the classifier is to attribute the reference to the head and mark it as more permanent, it will be closest to the head.

(1)

Nominal Group	Extreme	Internal	Thermodynamic	Conditions
Experiential Structure	Epithet	Epithet	Classifier	Thing
Logical Structure		Premodifier		Head
Type of Modification (linear)	δ	γ	β	α

Example (1) shows that the noun *condition* serves as the semantic object, or the Thing, and is the head of the nominal group modified by three premodifiers (*extreme, internal and thermodynamic*). However, the three premodifiers do not exhibit the same features because *extreme* and *internal* are (central) adjectives and they function as epithets in the given nominal group, while *thermodynamic* is a more peripheral adjective with the function of a classifier. The logical structure of the nominal group reveals the order of the premodifiers. This shows that the logical component of the nominal group is recursive—working leftwards from the head, thus raising the following questions: *What kind of conditions? What kind of thermodynamic conditions? What kind of internal thermodynamic conditions?* The questions that are raised here are concerned with the degree of acceptability of structures such as **internal extreme thermodynamic conditions,* and the correct sequence of modifiers within a nominal group when they belong to the same word class. There is a wide discussion on this matter to be found across literary sources (cf. Quirk, Greenbaum, Leech & Svartvik, 1985: 437). One of the conclusions is that the position of premodifiers in the nominal group is not completely free, but will depend on the meaning and type of premodifier (Pastor Gómez, 2010). The logical structure of the nominal group is equally important in understanding the order of the premodifiers or whether there exists a linear modification relationship. This is also referred to as a *stacked modification* (Payne & Huddleston, 2002: 446–7), as in (2a), or *non-linear modification (submodification)* between the components of the structure, as in (2b). In a stacked modification, all of the components modify the head successively, whereas in submodification there is a simultaneous modification of elements at a word level (i.e. below the group level because groups shift rank and behave like words in this case).

(2) a.

Nominal Group	Self-priming	Piston	Pump
Logical Structure	Premodifier	Premodifier	Head
Type of Modification (linear)	γ	β	α

'a pump that has a piston to force liquid from the inlet side to the outlet side and which is capable of evacuating air from the pump suction line without any external auxiliary device'

(2) b.

Nominal Group	Exhaust	Valve	Wear
Logical Structure	Premodifier	Head	Premodifier
Type of Modification (linear)	ββ	βα	α

'the wear of a valve that is used for releasing burnt gases from the cylinder'

In (2a), the structure of the nominal group is recursive, which can be verified by asking the following questions: *What kind of a pump? What kind of a piston pump?* In (2b), the same questions would not yield meaningful responses *(What kind of wear? - *exhaust wear)* since this nominal group represents the contraction of the underlying clause structure in which the noun *valve* functions as the object and the modifier specifies the type of object. *Exhaust* and *valve* form an individual nominal group, which shifts rank to word level and behaves like a single word. Both nouns *(exhaust* and *valve)* simultaneously modify the head *wear* of the complex nominal group *exhaust valve wear*. Hence, this nominal group cannot be interpreted as *exhaust wear; however, interpretations such as *valve wear/wear of the exhaust valve* are possible.

When more than two nouns premodify the head, the latter modifier is the object (e.g. *aluminium* in (3)) of the underlying sentence (*the alloy contains aluminium*), while the former noun (*strength*) designates a means, material, space, and so on (Quirk, Greenbaum, Leech & Svartvik, 1985: 1342). Epithets such as *high* give additional meanings to the noun:

high-strength aluminium alloy – *aluminium high-strength alloy (The alloy contains aluminium.)

Pastor Gómez (2010) argues that the number of items that may appear in premodifying position is theoretically unlimited. Nevertheless, finding more than four is unusual because "too much complexity in NP modification will imply a processing overload, leading to a loss of meaning and content" (Pastor Gómez, 2010: 10). This can be illustrated by example (4), in which the content load of the nominal group is too 'heavy' because it consists of 11 components. Hence, identifying the head or type of premodification, as well as the correct interpretation of the complex nominal group, becomes almost impossible.

(4) *Internal Combustion Auto Engine Piston Car Pendant Alloy Keychain Key Ring*

To be able to understand technical language, we need to look beyond vocabulary and shift our focus on grammar, or rather the interaction between grammar and vocabulary (Halliday, 1993). Pastor Gómez (2010: 21) further argues that "shared knowledge between interlocutors causes a reduction in explicitness"; that is, speakers do not repeat what is already known or can be inferred from previous utterances, or, in Quirk's words, "linguistic interchanges" (cf. Quirk, Greenbaum, Leech & Svartvik, 1985: 1234). Thus, efficient communication implies that the least number of words, groups, and/or clauses is used to express a given content. However, the ESP teacher does not necessarily possess this 'shared knowledge' owing to the fact that he/she is not an expert in the specialised field and needs to rely on linguistic means to interpret complex nominal groups, such as the structure of the nominal group itself. Therefore, the main research question is: can the linguistic structure of the nominal group be helpful in understanding the extralinguistic content? Given that terms are universal concepts of a specific domain, information on syntax and terminology knowledge may help ESP teachers to understand the extralinguistic content and enable them to teach complex nominal groups more successfully.

3. CORPUS AND METHODOLOGY

The following study is corpus-based (cf. Tognini-Bonelli, 2001), which means that corpus data is used "to support intuitive knowledge, to verify expectations, to allow linguistic phenomena to be quantified, and to find proof for existing theories or to retrieve illustrative samples" (Storjohann, 2005: 9). Being corpus-based in nature, the study analyses corpus data in order to ascertain if there is any connection between the number of frequencies of a nominal group and the position it occupies within another (complex) nominal group; that is, of a head or a premodifier.

Corpus compilation and term extraction were carried out as follows. A corpus of 59,565 tokens (referred to in this paper as Marine Engineering articles, MEA) was compiled from scientific articles to determine which nouns are most frequently used in the marine engineering scientific discourse (see Appendix 1). Another corpus counting 369,274 tokens (henceforth, Marine Engineering books, MEB), was compiled from five of the most relevant books for marine engineers according to *Marine Insight,* with the purpose of ascertaining the most frequent nouns in university and student books of marine engineering. The list of most frequently used nouns from this corpus can be found in Appendix 2.

The comparison of the two frequency lists (Appendices 1 and 2) yielded the following seven nouns as the most relevant (and common) nouns in the specialised field under study: *engine, fuel, diesel, cylinder, combustion, valve,* and *pump.*

The identified nouns were cross-referenced in the special field terminology database Struna to establish their *termhood*. The concept of termhood is understood in its broadest sense here indicating any degree to which a stable lexical unit is related to domain-specific concepts (Wong, 2009). Even though the Struna database was primarily intended to store and terminographically manage standardised and harmonised Croatian terms from various subject fields, it contains their equivalents in English and, depending on the domain, several other European languages. A pragmatic view of termhood (Kageura & Umino, 1996) is adopted in the database, whereby termhood is determined by the specialist of the field as well as the terminologist. Furthermore, we found words such as *piston, crankshaft, camshaft, gear* and *lube* to also be relevant in the marine engineering domain, basing our decision on a 10-year-long experience in teaching marine engineering English, as well as following advice from colleagues and terminologists whose efforts have been dedicated to standardising the marine engineering terminology. Accordingly, the 12 words were added to the list of domain-relevant seed words for the purposes of the study's main corpus compilation (henceforth, Marine Engineering_web, MEW) counting 1,257,782 tokens with the corpus compilation being carried out in Sketch Engine. After collecting the desired language sample, we used the Sketch Engine terminology extraction feature to extract terms from the main corpus. Automatic term extraction (ATE), as performed in Sketch Engine, is a hybrid model that combines both statistical and linguistic models by relying both on pattern matching and frequency count as well as linguistic criteria (cf. Sketch Engine, Quick Start Guide, 2019). The statistical model extracts terms by comparing the frequencies of the specialised corpus—also referred to as the focus corpus (in this case, MEW)—against a general language corpus—which is also called a reference corpus (in this case, English Web 2013 or the English Web 2013 sample)—and by calculating the keyness score (cf. Kilgarriff, 2009).5 The results are: 1) keywords or single-token items that appear more frequently in the focus corpus than in the reference corpus, and 2) terms or multi-words consisting of two or more tokens that appear more frequently in the focus corpus than in the reference corpus and at the same time are formatted as a term in the language, with the format being defined by the term grammar of a specific language (cf. Sketch Engine, User Guide, 2019). The latter feature—Term Structure—matches the potential structure of the term in the language (in our case, N1 + N2 + N3 + NN; and Adj+ N1 + N2 + NN, or any combination of the two), which is made possible through linguistic term extraction tools such as POS tagging and lemmatisation.

Due to terminological discrepancies employed by the Sketch Engine tool whereby single or simple terms are designated as keywords and multi-terms are designated as terms, for reasons of clarification we shall henceforth use 'simple term' to designate a single term or one token unit and 'complex term' to designate (multi-) terms; i.e., two or more token units.

The ATE feature from Sketch Engine yielded 986 complex terms (two or three token units) with the highest keyness score of 560.460 for *cylinder head* (the results for the

synonymous terms *lube oil, lubricating oil* and *lubrication oil* are presented in total) and the lowest keyness score of 11.710 for *shaft* axis. Table 1 shows the keyness score of the five most frequent complex terms from MEW proving they are terms from the field under study.

Term	Tag	Frequency (MEW)	Frequency (English Web 2013 Sample)	Keyness Score
Lubrication oil, lube oil, lubricating oil	[lemma="lubricating|lubricating|lube"] [lemma="oil"]	1,544	21	1244.040
Cylinder head	[lemma="cylinder"] [lemma="head"]	723	34	522.180
Connecting rod	[word="connecting"] [lemma="rodl"]	578	4	469.840
Fuel oil	[lemma="fuel"] [lemma="oil"]	664	103	382.580
Internal Combustion Engine	[lemma="internal"] [lemma="combustion"] [lemma="engine"]	427	79	264.950

Table 1. Frequency and keyness score of complex terms extracted from MEW.

Obtaining complex terms or nominal groups from the corpus cannot be carried out directly, which means that we cannot query the corpus for all premodifiers, heads, and postmodifiers, but instead have to direct the search for lemmas or POS.

As was discussed in Section 2, a nominal group consists of a head, which is usually a noun that can be premodified by adjectives, participles, and also other nouns, and postmodified by nouns, prepositional phrases, and relative clauses. Other types of postmodifiers, such as adjectives, are also possible (cf. Borucinsky, 2015).

In the attempt to support the research question with optimal data, the following CQL (Corpus query language) was used. In order to find nominal groups (complex terms) occupying the head position within another (complex) nominal group, the following query was used:

[tag="N.*|J.*"]{1,2}[lemma="lubricating|lubrication|lube"][lemma="oil"][!tag="N.*"]

This can be interpreted as: find all instances of the lemmas 'lubricating|lubrication|lube' and 'oil' preceded by one or two tokens {1,2}, which can be either a noun (N.* indicates any noun) or an adjective (J.* indicates any adjective), after which there cannot follow another noun (! is used for exclusion). This means that the searched term can be followed by any other part of speech and any other token (.,:; etc.). This query resulted in (simple) nominal groups occupying the head position within a nominal group.

When querying the corpus for the same term as the premodifier, the following query was used:

[tag="N.*|J.*"]{1,2}[lemma="lubricating|lubrication|lube"][lemma="oil"][tag="N.*"]

where the parameters are the same as in the previous search but the right-hand context is set differently—the lemmas can only be followed by another noun. Since only the premodifiers and heads, and primarily nominal groups consisting of nouns and adjectives are analysed, prepositional phrases and clauses as postmodifiers were not accounted for in the search. For example, a nominal group with a head *oil pump* is followed by another noun in 36 instances and by a preposition in 14 instances.

Instead of a lemma, a word search was used in order to include terms such as *lubricating* which can be both a participle and a noun, given the fact that POS tagging is not 100 % accurate. A word search was used in all instances where we believed that lemma search would not yield relevant results. For instance, the search for *lubricating oil* where *lubricating* is a lemma would yield 448 instances of heads of nominal groups and 110 instances of premodifiers, whereas the search for *lubricating* as a word would yield 498 instances of heads and 178 instances of premodifier.

Another important feature to decide upon was how many nouns and/or adjectives to account for in the premodifying, and how many nouns in the postmodifying position. Following a number of attempts, optimal results were achieved with one or two nouns and/or adjectives in premodifying position and one noun in postmodifying position. Terms with multiple spelling, such as *crank pin* and *crankpin*, were found to be somewhat problematic and in such cases both instances were accepted as heads and duplicate information was removed.

4. RESULTS AND DISCUSSION

Once the CQL parameters have been set, 200 nominal groups were analysed to determine whether they occupy the position of a head or a premodifier in a complex nominal group and whether there is a correlation between that position and the keyness score. The results are illustrated in Figure 1.
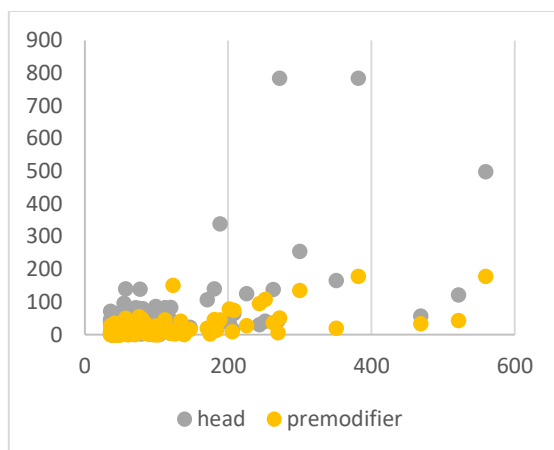
Figure 1. Distribution of head/premodifier position.

Figure 1 shows that there is no correlation between the position (head or premodifier) that a nominal group occupies within another (complex) nominal group and its keyness score or frequency in general, which proves that corpus frequency alone cannot help the ESP teacher disambiguate and better understand nominal groups.

The next step included analysing concordances of the most frequent complex nominal groups shown in Table 1, which also contained one of the 12 keywords listed in Section 3 (except for the term *connecting rod*). These were analysed to determine their structure: first, type (Type 1: N1 + N2 + N3 + NN; Type 2: Adj + N1 + N2 + NN or any combination of the two); second, number of constituents (3, 4, 5 or more); third, position within a nominal group (head or premodifier); and, finally, type of modification (linear or non-linear). The results are shown in Tables 2–6. The results do not indicate frequency counts or percentages but are merely an illustration of different types of nominal groups found in the corpus.

4.1. The nominal group containing the term variants *lube oil, lubrication oil, lubricating oil*

Nominal Group	Type	No. of Const.	Head or Premodifier	Type of Modification
Auxiliary engine lube oil	Type 2	4	Head	Non-linear
Cylinder lube oil/cylinder lubrication oil/cylinder lubricating oil	Type 1/Type 2	3	Head	Linear
Lubrication oil system	Type 1	3	Premodifier	Linear

Engine lubrication oil circuit	Type 1	4	Premodifier	Non-linear
Turbocharger lube oil storage tank	Type 1	5	Premodifier	Linear and non-linear

Table 2. Structure of nominal groups containing term variants: *lube oil, lubrication oil, lubricating oil.*

4.2. The nominal group containing the term *cylinder head*

Nominal Group	Type	No. of Const.	Head or Premodifier	Type of Modification
Individual cylinder head	Type 2	3	Head	Linear
Aluminium alloy cylinder head	Type 1	4	Head	Linear
Dual intake cylinder head port	Type 2	5	Premodifier	Linear and non-linear
Cylinder head temperature	Type 1	3	Premodifier	Non-linear
Cylinder head mating face	Type 1	4	Premodifier	Linear and non-linear

Table 3. Structure of nominal groups containing the term *cylinder head.*

4.3. The nominal group containing the term *connecting rod*

Nominal Group	Type	No. of Const.	Head or Premodifier	Type of Modification
Phosphatized bushingless connecting rod	Type 2	4	Head	Linear
Connecting rod assembly	Type 2	3	Premodifier	Non-linear
Connecting rod small end	Type 2	4	Premodifier	Non-linear
Connecting needle bearing	Type 2	4	Premodifier	Non-linear
Connecting rod angular speed radians	Type 2	5	Premodifier	Non-linear

Table 4. Structure of nominal groups containing the term *connecting rod.*

4.4. The nominal group containing the term *fuel oil*

Nominal Group	Type	No. of Const.	Head or Premodifier	Type of Modification
Heavy fuel oil	Type 2	3	Head	Linear
High pressure fuel oil tubes	Type 2	5	Premodifier	Non-linear
Fuel oil pump	Type 1	3	Premodifier	Linear
Fuel oil supply position	Type 1	4	Premodifier	Non-linear

Table 5. Structure of nominal groups containing the term *fuel oil.*

4.5. The nominal group containing the term *internal combustion engine*

Nominal Group	Type	No. of Const.	Head or Premodifier	Type of Modification (if applicable)
Reciprocating internal combustion engines	Type 2	4	Head	
Liquid fuelled internal combustion engine	Type 2	5	Head	
Hydrogen powered internal combustion engine	Type 2	5	Head	
Internal combustion engine components	Type 2	4	Premodifier	
Internal combustion chamber engine cycle	Type 2	5	Premodifier	

Table 6. Structure of nominal groups containing the term *internal combustion engine.*

Firstly, both types of structure (N1 + N2 + N3 + NN; Adj + N1 + N2 + NN) are represented in the corpus, even though the latter type will be easier to understand because adjectives commonly assume the function of epithets and rarely that of classifier. The type 1 structure can be as complex as 5 constituents, whereby the first or the last two nouns (e.g. *connecting rod*) form a fixed unit, shift rank from group to word level, and function as modifiers of the head in the complex nominal group. The head in this case is most commonly

a general noun; that is, a noun which does not necessarily belong to the specialised field, such as *capacity*, *quality*, or *consumption*, and so on or a very concrete noun, such as *spider, bolt, cage,* or *stud*. This structure of the nominal group is non-linear, as can be seen from example (5).

(5)

Nominal Group	Connecting	Rod	Assemby
Experiential Structure	Classifier	Thing	Thing
Logical Structure	Premodifier	Head	
	Premodifier		Head
Type of Modification (non-linear)	ββ	βα	α

However, when the analysed complex term occupies the head position in a nominal group and is modified by an adjective (e.g. *high strength connecting rod*), there is also a linear modification.

(6)

Nominal Group	High	Strength	Connecting	Rod
Experiential Structure	Epithet	Classifier	Classifier	Thing
Logical Structure	Premodifier	Premodifier	Head	
Type of Modification (linear and non-linear)	γ	β	α	

Second, an unexpected number of four or five constituent nominal groups was encountered in the corpus, despite the fact that such nominal groups are more difficult to process and understand. Example (7) is ambiguous and can be interpreted in two ways:

(7) a.

Nominal Group	Cylinder	Head	Mating	Face
Experiential Structure	Classifier	Thing	Classifier	Thing
Logical Structure	Premodifier	Head	Premodifier	Head
	β	α	β	α
	Premodifier		Head	
Type of Modification (non-linear)	β		α	

(7) b.

Nominal Group	Cylinder	Head	Mating	Face
Experiential Structure	Classifier	Classifier	Classifier	Thing
Logical Structure	Premodifier	Premodifier	Premodifier	Head
Type of Modification (linear)	δ	γ	β	α

N + N structures, and even more so N + N + N structures, achieve compactness, which explains why nominal modifiers are so frequently used (Varantola, 1993). Nonetheless, as can be seen from the previous examples, compactness is achieved at the expense of explicitness, which can lead to syntactic ambiguity of complex nominal groups because the meaning relations between the elements in a nominal group are affected. Of all the possible interpretations that an N1 + N2 + N3 + NN structure may have, speakers tend to choose the one that is most contextually plausible (Pastor Gómez, 2010) and when "an extralinguistic referent is in the air ambiguity fades away" (Pastor Gómez 2010: 66.). However, the ESP teacher may not be familiar with the extralinguistic referent and will thus have difficulty disambiguating and understanding the correct meaning of the nominal group.

Third, complex terms occupy both the position of a head and premodifier.

Fourth, the content load in nominal groups consisting of five nouns is very high. Nevertheless, processing them does not pose significant difficulty if the criteria for linear and non-linear modification are taken into account.

(8)

Nominal Group	Reciprocating	Internal	Combustion	Engine
Experiential Structure	Classifier	Epithet	Thing	Thing
	Premodifier	Premodifier	Head	Head
Logical Structure	Premodifier	Premodifier		Head
Type of Modification (linear and non-linear)	γ	β		α

'an engine which has a piston that reciprocates (moves up and down) and in which fuel is burned/combusted inside (internally)'

The following question can be raised here: Why is it not correct to say *internal reciprocating combustion engine?* According to Quirk, Greenbaum, Leech and Svartvik (1984) and the traditional grammar books of English, central adjectives such as *internal* should precede participles when both premodify the head of the nominal group. The answer lies in the complexity of the nominal group and type of modification. In case of linear modification, the adjective *internal* would precede the participle *reciprocating*. Nevertheless, if we consider the fact that the modification is not linear because the adjective *internal* modifies the head *combustion* and together they form the head of a separate nominal group and assume the function of premodifier in the complex nominal group, then it stands to reason that the participle precedes the central adjective in this nominal group. *Reciprocating* actually modifies the head *engine*, specifying that it has a piston that moves up and down and is not (for instance) a *jet engine*, so it is farther from the head than the nominal group *internal combustion*, which designates that the fuel is burnt inside the engine. The type of engine is specified first, and later the type of operation.

Let us consider a somewhat more complex example, such as the one presented in example (9).

(9)

Nominal Group	Turbocharger	Lube	Oil	Storage	Tank
Experiential Structure	Classifier	Classifier	Thing	Classifier	Thing
Logical Structure	Premodifier	Premodifier Premodifier	Head	Premodifier Head	Head
Type of Modification (linear and non-linear)	γ	β		α	

'a tank for storing the oil that is used for lubricating the turbocharger'

In example (9), there are two separate nominal groups (*lube oil* and *storage tank*), which shift rank to word level. Hence, it is inferred that the noun *storage* modifies the noun *tank*, and that the noun *lube* modifies the noun *oil*, and, finally, that the premodifiers *turbocharger* and *lube oil* mark the head *tank* as more permanent.

Our results seem to imply the advantageous role of SFG in the analyses of complex nominal groups and lend support to its active use in corpus searches. In addition, we have conducted informal interviews with non-ESP English teachers in order to infer their understanding of complex nominal groups. As expected, the interviewees encountered considerable difficulty in their attempts to process and interpret the given nominal compound since they were not familiar with the meaning or rather the function of particular engine components. Conversely, marine engineering students did not experience any difficulties in interpreting this nominal compound.

In spite of the informal character of the conducted interviews, they lead us to propose a major survey on how various groups—such as native speakers of English, English teachers (ESP and non-ESP), students, and specialists of the field—interpret complex nominal groups because it will help to find useful clues pointing towards the direction of how complex nominal groups are processed and interpreted.

5. CONCLUSIONS

The main motivation for carrying out this research laid in helping the ESP teacher overcome the fact that he or she is not a specialist of the field in question and how he/she can use insights gained from this research to better understand, and hence successfully teach, nominal compounds. The most relevant conclusion of the study is that in technical English, two-word

lexical units (e.g. *cylinder head*), and sometimes even three-word lexical units (e.g. *internal combustion engine*), behave as single units.

The corpus results have confirmed that complex nominal groups are a common feature of technical English. In this paper, we have proposed SFG as the method for the correct interpretation of complex nominal groups of the N1 + N2 + N3+ NN or Adj + N1 + N2 + N3 structure which can ultimately be employed by the ESP teacher who lacks the necessary extralinguistic knowledge to disambiguate between various senses. Complex terms consisting of two or three nouns become singularised—they shift rank and behave as lexical and not syntactic units. In a complex nominal group, they occupy either a (pre)modifier or a head position. The type of modification and the non-linear nature of the nominal group is the key to resolving syntactic ambiguity. Furthermore, examples which have been analysed using the SFG approach to grammar show the advantages of this theory over traditional approaches.

A study of how different language users—such as speakers of English who are not specialists in the field, English teachers (ESP and non-ESP), students and specialists of the field—interpret, understand and process nominal groups (i.e. which strategies they use for identifying constituents of a nominal group and how the complexity of the nominal group affects reading comprehension) would shed further light on this very complex and ambiguous topic.

NOTES

1 The term *noun phrase* is commonly used in linguistics (cf. Allerton, 1979; Chomsky, 1965; Gleason, 1961; Hudson, 1973; Lyons, 1968; etc.). In 1956, Halliday introduced the term *nominal group* and distinguished it from the term *noun phrase* (cf. Bloor & Bloor, 2004;

Halliday & Matthiessen, 1999). Given that this paper adopts the *Systemic Functional approach to Grammar* (SFG) method, the term nominal group will be used henceforth.

2 Cf. Johns and Dudley-Evans (1991).

3 On the distinction between modifiers and determinatives (cf. Zovko Dinković & Borucinsky, 2016).

4 This distinction is also referred to as reference-modifying and referent-modifying (cf. Bollinger, 1967).

5 The keyness score can be calculated using the following equation:

$$\frac{fpm_{rmfocus} + N}{fpm_{rmref} + N}$$

where $fpm_{rmfocus}$ is the normalised (per million) frequency of the word in the focus corpus, fpm_{rmref} is the normalised (per million) frequency of the word in the reference corpus, and N is the so-called smoothing parameter (N = 1 is the default value) (cf. Kilgarriff, 2009).

REFERENCES

Allerton, D. J. (1979). *Essentials of grammatical theory*. London: Routledge and Kegan Paul.

Biber, D., Stig J., Leech, G., Conrad S. & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education Limited.

Bloor, T. & Bloor, M. (2004). *The Functional Analysis of English, 2nd* ed. London: Edward Arnold.

Bollinger, D. (1967). Adjectives in English: Attribution and predication. *Lingua*, *18*, 1–34.

Borucinsky, M. (2015). *Modifikacija u imenskim skupinama u engleskome i hrvatskome jeziku*. Unpublished Doctoral Dissertation, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Fawcett, R. P. (2010). A Theory of Syntax for Systemic Functional Linguistics. *Current Issues in Linguistic Theory, 206.* Amsterdam & Philadelphia, PA: John Benjamins Publishing Company.

Fawcett, R. P. (2017). What is a system? What is a function? *Routledge Handbook of Systemic Functional linguistics*. doi: 10.4324/9781315413891.ch3

Fletcher, W. H. (2011). Corpus Analysis of the World Wide Web. In C. A. Chapelle (Ed.), *Encyclopedia of Applied Linguistics.* Wiley-Blackwell. Retrieved October 10, 2018 from http://www.encyclopediaofappliedlinguistics.com

Gleason, H. A. (1961). *An introduction to descriptive linguistics, rev.* ed. London: Holt Rinehart and Winson.

Halliday, M. A. K. & Matthiessen, C. M. I. M. (1999). *Construing Experience through Meaning: A Language-Based Approach to Cognition.* London & New York, NY: Cassel.

Halliday, M. A. K. & Matthiessen, C. M. I. M. (2004). *An Introduction to Functional Grammar, 3rd* ed. London: Edward Arnold.

Halliday, M. A. K. (1956). Grammatical categories in modern Chinese. *Transactions of the Philological Society,* 177–224.

Halliday, M. A. K. (1978). Language as Social Semiotic: *The Social Interpretation of Language and Meaning*. London: Edward Arnold.

Halliday, M. A. K. (1985). *An Introduction to Functional Grammar,* 1st ed. London: Edward Arnold.

Halliday, M. A. K. (1993). Some grammatical problems in scientific English. In M. A. K. Halliday & J. R. Martin (Eds.), *Writing Science* (pp. 69–85). London: The Falmer Press.

Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*, 2nd ed. London: Edward Arnold.

Hudson, R. A. (1973). *English complex sentences: An introduction to systemic grammar*. London: North Holland Publishing Co.

Johns, A. M. & Dudley – Evans, T. (1991). English for Specific Purposes: International in Scope, Specific in Purpose. *TESOL Quarterly*, 25, 2. doi: 10.2307/3587465.

Justeson, J. S. & Katz, S. M. (1995). Technical terminology: some linguistic properties and an algorithm for identifcation in text. *Natural Language Engineering*, *1*, 9-27. doi:10.1017/S1351324900000048

Kereković, S. (2016). Višeznačnost u jeziku struke. In S. L. Udier & K. Cergol Kovačević (Eds.), *Metodologija i primjena lingvističkih istraživanja* (pp 219-236). Zagreb: Filozoski fakultet.

Kilgarriff, A. (2009). Simple maths for keywords. In M. Mahlberg, V. González-Díaz & C. Smith (Eds.), *Proceedings of Corpus Linguistics Conference CL2009*. University of Liverpool.

Kozerenko, E. (2004). *A Hybrid Model for Language Structures Disambiguation in Machine Translation*. Moscow: Institute for Informatics Problems of the Russian Academy of Sciences. Retrieved March 12, 2018 from https://pdfs.semanticscholar.org/c6ee/ab8af0b1af6999228649503a970f01dd4e1c.pdf

Lyons, J. (1968). An introduction to theoretical linguistics. Cambridge: Cambridge University Press.

Martin, J. R., Matthiessen, C. M. I. M. & Painter, C. (2010). Deploying functional grammar. Beijing: The Commercial Press.

Pastor Gómez, I. (2010). Nominal Modifiers in Noun Phrase Structure: Evidence from Contemporary English. Unpublished Doctoral Dissertation, University Santiago de Compostela, Spain.

Payne, J. & Huddleston, R. (2002). Nouns and noun phrases. In R. Huddleston & G. Pullum (Eds.), The Cambridge Grammar of the English language (pp 323-524). Cambridge: Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). A comprehensive grammar of the English language. London: Longman.

Sketch Engine User Guide. Retrieved 1 April, 2019 from https://www.sketchengine.eu/user-guide/user-manual/term-extraction/

Storjohann, P. (2005). Corpus-driven vs. corpus-based approach to the study of relational patterns. In Proceedings of the Corpus Linguistics Conference 2005. Birmingham: University of Birmingham. Retrieved April 5, 2019 from https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/5006/file/Storjohann_Corpus_driven_vs_corpus_based_approach_to_the_study_of_relational_patterns_2005.pdf

Tognini-Bonelli, E. (2001). Corpus Linguistics at Work. Amsterdam/Philadelphia, MA: Benjamins.

Varantola, K. (1993). Modification of nouns by nouns. Bad by definition? In A. H. Jucker (Ed.), The Noun Phrase in English. Its Structure and Variability (pp 69-83). Heidelberg.

Warren, B. (2003). The role of links and/or qualia in head-modifier construction. In B. Nerlich, Z. Todd & V. Herman (Eds.), Polysemy: Patterns of Meaning in Mind and Language (pp 233-252). Amsterdam and Philadelphia, MA: John Benjamins.

Wong, W. Y., Liu, W. & Bennamoun, M. (2009). Determination of Unithood and Termhood for Term Recognition. In M. Song, & Y. B. Wu (Eds.), *Handbook of Research on Text and Web Mining Technologies* (Vol. 2, pp. 500-529). USA: IGI Global. doi: 10.4018/978-1-59904-990-8.ch030

Zhoulin, R. (2016). A Corpus-Based Functional Analysis of Complex Nominal Groups in Written Business Discourse: The Case of "Business". International Journal of Computer-Assisted Language Learning and Teaching, 6/ 2, 74-90.

Zovko Dinković, I. & Borucinsky, M. (2016). On determiners in Croatian from the perspective of Systemic Functional Grammar. Jezikoslovlje, 17/1-2, 7-22.

Web sources:

Marine Insight - https://www.marineinsight.com/

Sketch engine start guide - https://www.sketchengine.eu/quick-start-guide/

Struna - http://struna.ihjj.hr/

APPENDIX 1: FREQUENCY LIST FROM MARINE ENGINEERING_ARTICLES (59 565 TOKENS)

Item	Freq
engine	683
fuel	339
system	315
engines	257
gas	247
exhaust	225
heat	215
diesel	211
propulsion	198
power	187
temperature	165
cylinder	156
model	154
combustion	150
marine	145
figure	144

parameters	123
operation	113
ship	109
pressure	106

APPENDIX 2: FREQUENCY LIST FROM MARINE ENGINEERING_BOOKS (369 274 TOKENS)

Item	Freq
figure	1679
water	1404
air	1220
system	1214
pressure	1153
oil	1103
valve	1066
pump	1063
control	890
ship	860
cargo	686
deck	664
steam	660
tank	636
type	629
shaft	569
temperature	562
fuel	561
ships	556
systems	547