

“Phrasal compounds” and the discourse/lexicology interface: “conglomeration” within the French tradition of English lexicology

Nicolas Ballier, Université Paris Diderot

Abstract

*This paper tries to illustrate the relevance of what Benveniste called *conglomérés* (conglomerates) for the study of English. It shows that conglomerates refer to a set of regular phenomena and account for two general trends of the lexicon : the quest for economy on the one hand and the importance of speech (*parole*) in the lexicalisation of units on the other hand. By analysing a borderline case of morphology in two senses of the word (morphology vs. syntax and phrases vs. compounds), a case is made for the investigation of a delocutive matrix, the cline between discourse and the lexicon. In that sense, this paper offers a partial review of French approaches to English lexicology within the tradition of enunciative linguistics.*

Keywords: *conglomerates, phrasal compounds, Benveniste, syntax/lexicon interface.*

1 Introduction

« Conglomeration » is the term used by Sebeok (1972 : 94) to describe the formation of “congloméré”, i.e. the word-formation process suggested by Benveniste (1966) to account for nominal lexemes in French of the type *va-nu-pieds*, *monte-en-l'air*, *décrochez-moi-ça*. In English, constructions such as *a forget-me-not* now tend to be referred to as “phrasal lexemes” as in Masini (2009), but this paper contends that this somewhat overlooked tradition is worth investigating. Benveniste’s still untranslated paper “Some new forms of nominal compounding”, from the second volume of his *Problèmes de linguistique générale* offers an original and powerful analysis for compounds with implications for linguistic analysis that could be elaborated, especially for the interface between discourse and the lexicon. Specifically, a revaluation of the

French morphological tradition (Darmesteter, Benveniste, Guilbert) as to conglomerates leads to reconsider a delocutive addition to the “lexicogenic matrices” for the English language put forward by Tournier (1988) to account for neologisms/ word-formation in English. This paper tries to advocate the relevance of what Benveniste called *conglomérés* (conglomerates) for the study of English and, more generally, the importance of delocutive outputs. It shows that conglomerates refer to a set of regular phenomena and account for two general trends of the lexicon : the quest for economy on the one hand and the importance of speech (*parole*) in the lexicalisation of complex units on the other hand.

This paper³⁸ is an attempt at characterizing an under-researched notion put forward by Benveniste and a bird's eye view of the French tradition of « enunciative » linguistics as applied to English, adumbrating the potential contribution of this theoretical framework to word-formation. Section 1 details the examples selected by Benveniste for French and shows the existence of similar constructions in English. Section 2 extends the comparison to other languages where the notion has been explicitly applied. In the spirit of Prof. Štekauer's interest for compounds and universals, a contrastive analysis will show the usefulness of such constructions for English, French, Bambara, Italian, and Slovak. Section 3 offers a preliminary corpus-based investigation of hyphenation sequences in LPD and CELEX. Automatic queries of the CELEX database yielded 163 entries with two or more hyphens; these were then further analysed, looking for similarities in stress patterns and construction. The criterion of multi-word hyphenation is discussed in relation to the economy principle as well as the stipulation that signifiers in conglomerates are compact as in *a whodunit*. Section 4 discusses the institutionalisation of these sequences and the relevance of the term ‘quotational compound’ to refer to similar citational constructions. The constraints on institutionalisation lead to a revaluation of Jean Tournier's matrices of English word

³⁸ The gist of this paper develops one aspect of my research agenda delineated in my Habilitation thesis (Ballier 2004). A PhD (Hučka, in prep) jointly supervised with Prof Štekauer has resulted from the idea of investigating conglomeration in English. Thanks are due to Pierre Arnaud, Pavol Hučka, Véronique Pouillon, Vincent Renner, Salvador Valera and the anonymous reviewers for comments on an earlier version. The standard disclaimer applies as to the remaining flaws of the paper.

formation (Tournier 1988). This cornerstone of English morphology in France could now be challenged or at least nuanced : hybrid forms have emerged with the internet, and a specific case can be made for a “delocutive” matrix, whereby complex lexical units actually emerge out of speech or discourse. Section 5 elaborates on Benveniste’s utterance-based morphological tradition, making the claim that a “delocutive” matrix ought to be considered, in relation to delocutive verbs. A critical reevaluation of the twelve “matrices” chosen by Tournier for his inventory of word-formation processes (*néologie*) as analysed in his magnum opus (Tournier 1988).

2 Conglomerates with Benveniste and in the tradition of French morphology

This section presents the original passage from Benveniste 1966, delineates the two lexicalisation clines that can be deduced from Benveniste's examples and shows the vitality of the term in French terminology as applied to French.

2.1 Genesis of the term

« Conglomération » is the term used by Sebeok (1972 : 94) to translate the word-formation device suggested by Benveniste to account for terms such as *va-nu-pieds*, *monte-en-l'air*, *décrochez-moi-ça*³⁹. It is tempting to differentiate between the process (*conglomération*) and for the result (*conglomerate*). Conglomerates are described and distinguished from compounds in the following passage:

We will distinguish conglomerates from compounds. We apply this term to new units formed from complex phrases of more than two units. Some are predicative phrases that have been converted into substantives: thus : *va-nu-pieds* [tramp, beggar; literally *goes naked feet*], *meurt-de-faim* [lack-all, pauper; literally *dies of hunger*], *monte-en-l'air*

³⁹ I have contended in my PhD (Ballier 1997) that deverbal nouns used as metalinguistic terms benefited, sometimes ad nauseam from the aspectual processive ambiguities of its reference, sometimes allowing at times the linguist to kill two linguistic birds with one stone. Blends and blending corresponds in French to amalgames & amalgamation, the distinction between the result and the process is not so clear with (Fr.) amalgamation or conglomération. I enjoy the prospect of having conglomerate for the result and conglomération for the process.

[burglar; literally *goes up in the air*], *décrochez-moi-ça* [secondhand clothes store; literally *could you take down that [hanger] for me*]. Others are adverbial phrases where archaic elements survive. For instance, *dorénavant* [henceforth] (= *d'ore en avant* [i.e. *from now forward*]), *désormais* [as of now] (= *dès or mais* [i.e. *from right now*]) can no longer readily be analyzed, and *jamais* not at all; in *aujourd'hui* [today], however, one may still perceive the initial sequence « *au jour d' ...* » [on the day of], as the three parts of *auparavant* [previously], « *au par avant* » [at the by before], even if their syntactic organization is not immediately obvious. Similarly, the old predicative phrase *n'a guère* [not has much] is tightened into the more familiar *naguère* [a short while ago]. The common feature of these conglomerates is that a complex construction becomes fused into a single whole, without its constitutive elements being mutilated or altered. These may be completely or partially recognizable, depending on how far the conglomerate dates back: in *justaucorps* [jerkin] (« *juste au corps* » [tight to the body]) can be easily retrieved; in *gendarme* [policeman], a recourse to the plural form is necessary for *gens* to be understood [as the head of the phrase « *gens d'armes* », i.e. *armed men*]. Generally, conglomerates tend to be compact signs. [Benveniste 1966, our translation]

This definitional paragraph is the only mention of conglomerates in Benveniste's article, which mostly focuses on a form of compounding (involving technical terms) more often commented : *synapsies* (N prep N constructions of the type *machine à laver* (literally *machine for washing*, 'washing machine'). The following features are apparently selected by Benveniste : 1) conglomerates require three elements or more, a feature sometime neglected in the literature, where *rendez-vous*⁴⁰ is sometimes taken as

⁴⁰ We may object to *rendez-vous*, since the original utterance would not only include two words (lit. go you) but also prepositions phrases indicating time and place (see you at 4 in Paris). A multi-word unit superior to two words or more might be more consistent with a cross-linguistic analysis. *Rendez-vous* comes from a pronominal verb and has no argument. A two-word sequence means only one hyphen and little distinction with compounds. Coordinated structures such as *le va-et-vient* (to and fro movement, also used to refer to an electrical switch) are in line with the minimal structural characterization of the phrase as having at least three words. I have no space for such discussion but bigrams should also be discussed within idioms *va-tout* ; *va-vite* are hyphenated when the conglomerate could be said to be *jouer son va-tout* to stake (risk one's all) and *à la va-vite* ('quick and dirty'). The 64 occurrences in FRANTEXT of *va-vite* where for 'à la va-vite'. *Va-tout* had 56 occurrences, 2 for *faire* and the rest for *jouer* [possessive determiner] *va-tout*. For practical reasons, I favour trigrams for the investigation of

an example of a conglomerate, 2) they are converted predicative phrases, where the original order of constituents is retained, 3) two clines are considered: one is nominal and the other adverbial. The final feature, “compactness” (*gendarmes* for *gens d’armes*) does not necessarily correspond to a single cline, since *gendarmes* is nominal and *dorénavant* adverbial. I will elaborate on these three features in the next two sections.

2.2 Two lexicalisation clines in French?

Beside this nominal analysis of the argument structure, conglomeration seems to imply two clines in French, at least if we stick to Benveniste’s examples. The end-product of conglomeration favours adverbs on an adverbial cline (from PP to adverbs); we could add to Benveniste’s examples this archaic interjection *par le sang de Dieu* [by the blood of God] > *par le sang bleu* (by the blue blood, standard euphemism) > *palsambleu* (*blimey*). This conglomerate adverb tends to be opaque and does not imply a matrix predicate. The second cline concerns grammatically invariable nominal heads, with an output similar to a phraseological unit, as explained in Dubois & Dubois (1971). They note that phrases retain morphemes of tense, mode, person (-*ez* in *décrochez*) and also keep the order of the constituents of the initial utterance. For example, the nominal conglomerate *un suivez-moi-jeune-homme* (lit. *A follow me young man*, a kind of ribbon) is based on a matrix utterance (“*suivez-moi, jeune homme*” / *follow me, young man*) where the flectional morpheme (-*ez*) is retained. Comparing with French compounds, they observe that *porte* (*carries*) in *porte-bagages* (*rack*) is invariable. They explain the emergence of conglomerates over time by a nominalization of a sentence or sequence. Conversely, they note that adverbs resulting from a qualifying process (*naguère*) do not result from a predicate and observe that the compositionality of these constituents has become opaque.

conglomerates in English because trigrams are more efficient than bigrams for data-mining and because compounds would cause too much noise for queries based on single hyphenated sequences.

2.3 *The legacy of the term*

The term *conglomerate* has sometimes been alluded to in the analysis of English compounds (Boucher 1992), and its intellectual background was studied by Léon (2004). A working definition might be a complex polylexical unit involving more than one item. It has to be distinguished from compounds as it precludes recursivity. Rikus (1978:561) describes them as “fragments de syntagmes-événements” (fragments of event-phrases) and details the argument structures to show that adverbials, but not objects, are involved in the predication (*monte-en-l’air*). Indeed, conglomerates instantiate syntactic functions beside objects. Potential examples of conglomeration include *monte-en-l’air*, which exhibits interesting properties (on top of the minimal trigram with two successive hyphens) as compared with standard nominal compounds such as *monte-charge* (*a goods lift*). The underlying argument structure Verb-Object is instantiated in the compounds, whereas in *monte-en-l’air* only is the adjunct expressed. Similar analyses can be proposed for intransitive verbs such as *aller* (*va-de-l’avant*, *va-de-la-gueule*, *va-et-vient*, *va-nu-pieds*). Finally, Rikus (1978) makes the point that the distinction should be drawn with synapsies like *pomme de terre* (lit. ‘apple of earth’, *potato*). His conclusion in this section is similar to the point we wish to illustrate, the existence of a continuum between discourse and word-formation: « le langage oscille ainsi entre le discursif et le lexical, le second étant toujours tiré du premier, au moins quant au modèle sur lequel repose la formation analogique » (« language vacillates between discourse and lexicon, the latter being always drawn from the former, at least as to the model for analogical formations.”)

The concept is still vivid in the French morphological tradition (Fiala 1989, Mahrer 2011), though the resulting units tend to be analysed as the output of syntax. In her review paper of compounds in French, (Villoing 2012:35) rules out conglomerates (and synapsies), in the name of the syntax / lexicon divide. She gives examples of conglomerates (without using the term) to distinguish them from compounds, because they are “lexicalized phrases that behave like lexical units” (*le qu’en-dira-t-on*) or “lexicalized syntactic constructions that behave like lexical units”. Similarly, Fradin (2003) analyses conglomerates as sequences of units produced by syntax (*unités syntactiquement construites*). Fradin (2003: 202) distinguishes between *dorénavant*

(*unité logofigée*) and lexically constructed units such as *décrochez-moi-ça*. For the former, he makes the point that *malgré* and *beaucoup* may have had the adjective + noun format, but do not behave in syntax as a noun phrase (*malgré* is a preposition, *beaucoup* is a quantifier) since they are function words. For the latter, he explains (Fradin 2003:224) how these expressions are construed by speakers as the lexicalization of an utterance (“expressions lexicalisées par délocutivité”). He analyses the first kind as a form of grammaticalisation : the transformation of a sequence of lexical units into a lexical unit having a grammatical function such as adverb (*dorénavant*), connector (*puisque*) or quantifier (*beaucoup*).

Nowadays, conglomerates are typically referred to as "phrasal lexemes" in English (Masini 2009). The interconnection with discourse and syntax is being discussed. Masini 2009 calls them « phrasal lexemes » and analyse examples in Italian within a constructionist perspective. Similar constructions are also discussed in (Booij 2007), (Giegerich 2005) or (Booij 2002) for Dutch. Construction grammar appears as a particular well-designed framework to re-assess the boundaries between syntax and the lexicon and to tackle the “classification of constructions that populate the middle ground of the lexicon-syntax continuum and to the theoretical debate on the demarcation between words and phrases » (Masini 2009 : 256). Interestingly enough, when delineating phrasal compounds within the framework of construction grammar, Masini (2009) explicitly refers to the Benveniste 1966 paper, but to the other subtype of compounds she observes, namely *synapsies*, sequences of the type *N prep N*. This family of constructions labelled ‘phrasal lexemes’ focusses on a more regular syntactic type of constructions, but overlooks *conglomerates*. Unfortunately, this seminal paper « Formes nouvelles de la composition nominale » (New forms of noun compounding) originally published in the *Bulletin de la Société Linguistique* was never officially translated into English as it partakes to the second volume of *Problèmes de linguistique générale*. As a consequence, the term has rarely been used to analyse other languages than French, all the more so as the same paper also describes *synapsie*, a nominal compound type more often referred to in the literature.

In the wake of Benveniste’s work, it would be worth mentioning arguments in favour of other word-formation processes. Needless to say, in spite of some of the

shortcomings of the translation of the first volume, the second volume of his *Problèmes de linguistique générale* ought to be translated and his study on “New forms of nominal compounding” is worth being read even after forty years. Insights about *synapsies* and conglomerate forms are worth considering.

3 A quick cross-linguistic comparison

This section briefly examines the languages for which the notion has been explicitly applied in their description. In the spirit of Prof Štekauer’s interest for compounds and universals, a contrastive analysis will show the interest of such constructions for English, French, Bambara (data from Dumestre 2003 and Creissels 2004), Italian (data from Radimský 2006) and Slovak (data from the Slovak corpus). Formal features are put forward, such as the linear order of the constituents of the compound as part of the utterance in which they originally appear. More crucially, I show that univerbation is not a criterial property of conglomerates.

3.1 Bambara

In his grammar of Bambara, a Manding Language, Gérard Dumestre has a whole chapter dedicated to conglomerates. A footnote specifies that Benveniste’s term is somewhat adapted in the analysis, but comparable points are made (see also Creissels 2004), notably in relation to French. The typology⁴¹ he gives shows preference for nouns, though examples for verbs are given. He notes that, for a proportion too important to be due to chance, some nouns are familiar or humoristic. He gives a certain number of examples, such as *golo-be’-n’-s’en-na’ br* (litt. I have skin on my feet) a ‘policeman’. He has certain examples with lewd allusions such as the one encountered in French for *le baise-en-ville* (lit. ‘screws in city’, ‘overnight bag’). The order of the constituents of the initial utterance is preserved, as in the word for ‘gossip’ *fɔra’-n’-nyɛ’na* (*fɔra’* ‘was said’ + *n’* ‘me’+ *nyɛ’na* postposed loc.). He also gives several

⁴¹ Halaoui (1993) draws a subtle and interesting distinction between three types of conglomerates as used in Bambara.

examples of conglomerates that have become opaque: bɔ́nson yi ‘offspring’ < [bɔ́ -n’-s’ɔn] (to get out of me), mɛ́ndija dl ‘happiness’ < [n’-be’-n’-dija] (*I rejoice*).

3.2 Italian

Schwartz (2006) alludes to conglomerates when he discusses French and Italian examples under the heading ‘univerbation’. With univerbation comes the concept of a change in the head *pomo d’oro* has *pomi d’oro* as plural form, but *pomodoro* has *pomodori* as plural (Schwartz 2006 : 243). Formal properties for univerbation as listed in Schwartz (2006) include

- (1) several determiners are possible : *un m’as-tu-vu, une m’as-tu-vu*, (show, literally *have you seen me* ?)
- (2) pluralisation is possible (Schwartz even has *des matuvus*, which is mostly found as *m’as-tu-vus*)
- (3) lexemes can be suffixed : *le m’as-tu-vuisme*

Interestingly, *le je-m’en-foutisme, le m’as-tu-vuisme, le jusqu’au-boutisme* (*hardliners’ stance*) can be traced back to utterances such as ‘I don’t care’ (*je m’en fous*), and the concept of univerbation could be seen as the final stage of conglomeration. If *compacity* is taken as a criterion, then we still need a distinction between *va-nu-pieds* and *m’as-tu-vu*: only the latter could undergo dehyphenation and suffixation. *Pomo d’oro* would fall into the category of *synapsie* (typically within a prepositional phrase pattern, with a limited set of prepositions and no article in the PP,) univerbation describes the morphological process whereby the word is fused as a single unit: *pomodori*. Univerbation refers to a process describing a formal property of the resulting unit, but does not pre-conceive the morphological word-formation category. We should investigate the original construction before univerbation (dehyphenation) takes place. Un *matuvu* might be a univerbation, but it is a conglomerate, *va-nus-pieds* is a conglomerate, but not a univerbation and *pomodori* is a univerbation but not a conglomerate. Conglomerates include more complex structures than just prepositional phrases, they result more specifically from an utterance (note *ton je-m’en-foutisme*).

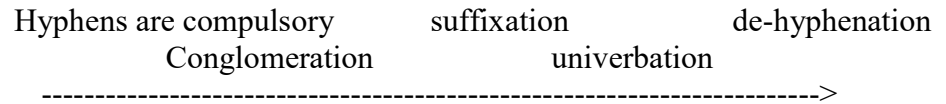


Figure 1 the univerbation / conglomerate cline

Even if economy is a driving force behind conglomerates, it seems to me that compacity is not a strict requirement for conglomerates, but rather a potential final stage in a cline, as illustrated in Figure 1. Similarly, Radimský (2006 :48) mentions *un saliscendi*, *un toccasana*, *un tiremmolla*, *un tira é molla*, *il va e vieni*, *il cessate-il-fuoco*. My contention would be that *il va e vieni* and *il cessate-il-fuoco* ought to be considered here as conglomerates. Though translation cannot be used as a criterion, as English is ambiguous with *cease-fire*, it should be noticed that French has a conglomerate for *le va-et-vient* and *un cessez-le feu*. Radimský (2006) makes the point that “lexicalisation of a part of the sentence” is too vague and fuzzy, as it may apply to *postacelere*, *climaterante* and Prep N compounds such as *senzacasa* (*homeless*). Again, maximum compacity is not a satisfactory criterion.

3.3. Slovak⁴²

For the sake of the argument, here are a couple of examples in Slovak that seem to fall with this range of constructions. Hyphenation is rare in Slovak but grammaticalisation of prepositional phrases as manner adverb can still be found without hyphens, for example *odušu* (lit. *for soul*, quickly, as if your soul depended on it) or *naverímboha* (lit. on I believe [in] God), which means *aimlessly*.

(4) všadebol (lit. *everywhere he was*, an excellent sportsman).

(5) bohvieкто (lit. *God knows who*)

⁴² I am indebted to Pavol Hučka for the following examples.

(4) and (5) are good candidates for such constructions. The noun *všadebol* can be used as a subject, predicatively and attributively with a surname. *Bohviekto* can be used as a pronoun expressing uncertainty as to the identity of the referent (very much like *God knows who* in English) and to refer to a person who thinks too highly of herself and can be used predicatively with a hint of criticism.

4 A dictionary-based investigation

We have seen that univervation is not a sufficient criterion; this section tests two of the putative properties of conglomerates for English and shows the limits of multi-hyphenation as a formal criterion to retrieve candidates for conglomeration as well as the irregularity of stress patterns - a question for those multi-word units already raised in Trevian (2003).

4.1 Multi-word hyphenation and conglomerates (a preliminary corpus-based study)

The corpus query is based on automatic extractions from *LPD* (Wells 1990) and the CELEX database (Baayen et al. 1995). The 163 hyphenated sequences found in the CELEX are analysed, looking for similarity in stress patterns and construction. The criterion of multi-word hyphenation is discussed in relation to the economy principle as well as the prediction / requirement that signifieds for conglomerates get compact as in *a whodunit*.

In a nutshell, hyphenated sequences can point to good candidates for conglomerates, as is the case with *might-have-beens* (both *CELEX* and *LPD*). The crucial distinction can be retrieved automatically: even though hyphenation can be chaotic in English, automatic queries can be programmed. For French, Labbé (1990) covers some of the lemmatisation issues in connection to the « agglutinating hyphens » which can be similarly observed in English for constructions such as *the what's-his-name*, *the what do you call it*, *the might-have-been*. In some cases as in (5), writers may abuse of this hyphenation device for playfulness:

(6) Work, the what's-its-name of the thingummy and the thing-um-a-bob of the what d'you-call-it. (In PG Wodehouse *Psmith, Journalist*; 1915)

At this stage, a very basic query of two databases allows the following observations. Queries for multiple hyphenated structures of the type $x - x_i - x_n$ of three items or more resulted in 164 hits in CELEX and 316 in LPD. The 164 multi-hyphenated sequences of the CELEX yielded sequences such as place-names (*Stoke-on-Trent, Wotton-under-Edge*), proper nouns (*Winnie-the-Pooh*), borrowings (*vol-au-vent*), 33 coordinated structures (*pepper-and-salt, pen-and-ink*) and good candidates for conglomeration such as *what-d'you-call-it, what-you-may-call-it, take-it-or-leave-it* or *over-the-counter*. LPD has more (and more exotic) place-names (*Beach-la-Mar, Dares-Salaam, Mont-Saint-Michel*) and interjections such as *fiddle-de-dee* (also spelt *fiddlededee*), 51 coordinated structures like *hide-and-seek, hit-and-miss, cloak-and-dagger* and other interesting predicates such as *holier-than-thou*. Comparing the pronoun *it* in conglomerates, CELEX has only *know-it-all*, whereas LPD also has *do-it-yourself, take-it-or-leave-it, what-d'you-call-it* and *what-you-may-call-it*. The different recalls in the two databases for the same structural hyphenated structures can partly be explained by the presence or graphic variants in LPD (24), but serve as evidence that hyphenation as such cannot be trusted as a necessary criterion for conglomerates. The most subtle example is *come-at-able* (that may be come at or reached, OED, unrevised entry), which can be found as *come-atable* or *comeatable*, though it is derivational.

4.2 Stress patterns

I like to fancy that initial or post-initial secondary stresses may serve in English as modulation of the signal for prosodic boundaries of a complex lexical unit (consider the famous criterion of early-stressed sequences sometimes put forward for compounds). Conglomerates may not support this view as data seems contradictory, as already evidenced in Trevian (2003). In a section dedicated to what he terms 'hyphenated multiword compounds', he opposes what he deems to be the vast majority of late-stressed compounds such as *cash-and-'carry* and the 198 exceptions he has spotted, among which *for'get me not*. He shows that allogenic units (borrowings from French)

show a remarkable regularity such as *je-ne sais-'quoi*, where *'cul de sac* is given as the only exception. For the hyphenated structures inventoried in Trevian (2003), primary stresses and not secondary stresses come into play, and they either operate on the initial item or on the final one. Dumestre (1993) mentions a specific tonal pattern for conglomerates in Bambara, but the evidence for English seems somewhat inconclusive for the time being, except that hyphenation is not a necessary condition for conglomerates. One of the reasons could be that multi-word hyphenation is sometimes presented as the signal of a quotation.

5 Utterance-based word-formation and “quotational insertions”

Section 4 discusses the institutionalisation of these sequences and the term ‘quotational compound’, which is used to refer to similar citational constructions and investigates the status of «quotational insertions» (Meibauer 2007).

5.1 Connection to quotational compounds

From that point of view, it is tempting to compare this kind of lexical units with quotational compounds. Dialogue-repeated sections - a recurrent device in Kathy Reich’s novels - is a case in point : as (i) *Dorval’s “welcome to Canada” immigration line usually makes* (ii) *Disney World’s snake-back-and-forth-through-the-ribbon-maze queue look short*. [*Bones to Ashes*, 156, quoted in Ryšavá 2012]. In her MA thesis, Ryšavá (2012), insists on the recontextualisation of an initial utterance.

A quotational compound is an expression that is cut out of its original environment and as a whole used in a different syntactic position, e.g. *I will take this food away with me. – take-away food*, etc. The origin of quotational compounds may range from short phrases (*face-to-face conversation* to finite clauses (*all-you-can eat menu, the I-didn’t-do-it look in her face*). [... T]he expressions cannot be strictly divided into two (or more) generally described groups. They are on a scale, where on one pole are these just described expressions, and towards the other pole there are expressions that were repeated over time, and therefore have become fixed and stable in the language, e.g. *a merry-go-round, a drive-through store* (Ryšavá 2012: 2).

Kathy Reichs often resorts to such constructions, sometimes mixing French and English, as evidenced in the following examples from the same novel:

- (7) Arnoldo's parts aren't zip-a-dee-doo-dah. (111)
- (8) Despite the "don't worry, be happy" attire, Hippo did not appear to be having a good day. (126)
- (9) My left hand did an automatic hair-behind-the-ears tuck, then I realized Ryan's remark was directed at the skull. (89)
- (10) Pierre LaManche is a large man in a grandpa-was-a-lumberjack sort of way. (29)
- (11) And oh-so-very-thoroughly très French (339)

Similarly, Nosek (1985, 159) defines these expressions like this "Quotational compounds are multiword groups such as *hand-in-glove* or *bride-to-be*". This gradient from discourse to lexical unit has to be investigated, but 'quotational compound' may not be the ideal term.

5.2 *Quotational compounds and the institutionalization of conglomerates*

I wish to argue against the term "quotational compound" for two kinds of reasons. First, it limits the scope of constructions to nominal compounds where the modifier is a quote. Secondly, this term is more accurate to refer to compounds whose interpretation depends on contextual information. I would prefer to reserve the term "quotational compounds" to compounds whose reference can only be elucidated in a given context. I understand *quotational compounds* as compounds, referring to a predicate where the identification of the referents of the NPs depend on left context. I exemplify this with a set of characters from Ruth Rendell, *Simisola* 235-237. In this passage, referents of NPs are resumed by compounds anaphorically referring to an initial predicate.

The spotty boy [...] was chewing the cuticles round his fingernails. His opposite neighbour with the pale dinosaur on his chest, just as Wexford approached, hit on the diverting idea of throwing pieces of gravel, of which he had a handful, at the stack of cans [...]

'Where's your friend?' [...]

No one answered. The smoker smoked, concentrating as if it was a study he was engaged in, involving memory and even powers of deduction. The cuticle-biter bit his cuticles and made more rings with his toes in the smoker's ash. The stone-thrower threw his handful of gravel over his shoulder and produced a packet from which he took a cigarette. (*Simisola* 235-236)

The point I would like to make here is that “*The spotty boy*” does not pose any problem for interpretation, even without this passage, but *the cuticle-biter* can only be understood in this context. The interpretation of these compounds is heavily context-dependent, it is a definite (unambiguously anaphoric) yet not specific reference, hence the importance of the (anaphoric) deictic use of *the* here. What a lexical unit as *bride-to-be* does is semantically different. It is not dependent on a context for interpretation, it is potentially less specific and not necessarily definite in its uses. The interpretation of conglomerates is not constrained as in nonce-formations by the initial quote or by anaphoric reference, *le qu'en dira-t-on* (what Mrs Grumpy will say), *le je-ne-sais-quoi* (*I don't know what*) partake of definite reference but is less specific. The benefit of my terminological twist is that in this case, “quotational compounds” are really compounds, whereas conglomerates present a wider spectrum of constructions.

Trips (2014) has ten sub-types of phrasal compounds where the first element is a quote such as a “*work or starve*” *philosophy* or a “*Weather hot, cricket wonderful*” *postcard* where she respectively describes the first element as “conjoined verbs” and “sentence with elided verb”). Her typology based on findings from the BNC allows the recognition of ten sub-types of phrasal compounds. I would like to point out that the set of examples does not encompass all the syntactic slots that conglomerates may occupy. Data is not limited to attributive positions to nouns, or just N1 specification of N1N2. I believe that conglomerates can occur in other syntactic slots. More generally, a *whodunit* (<*who has done it*) is distinct from NN compounding. Interestingly, some of the constructions that would fall into her categorisation of compounds may be used as a

noun of its own. An easy example is *over-the-counter* drug, which can be found in predicative uses, nominal uses, potentially pluralised and initialised (an OTC).

6 A revaluation of Tournier's word-formation matrices

An unsuccessful EU Socrates bid with Prof Štekauer had contemplated the possibility of discussing the breeding grounds and theoretical tenets in linguistics for the European PhDs of our partners in the project. Part of this section delineates the mainstream approach to English morphology as taught (if at all) in French universities as I see it.

6.1 *The general presentation of the lexicogenic matrices*

English morphology is not one of French linguists' strong points for English (see Ballier 1997 for an overview of the French theoretical underpinning of the enunciative school and some of the pet topics of grammatical research for English in France). In the 1990s, on a professional mailing list, Jean Tournier would refer to besieged and minority lexicologists of English dominated by grammarians as being "lexikosovars", a pun quite characteristic of his style but variously appreciated in the research community during Sarajevo's bombings. This subsection delineates the matrices of word-formation as presented in France's *magnum opus* of English lexicology in the eighties. Tournier's *Introduction descriptive à la lexicogénétique de l'anglais contemporain* published in 1985 was the revised version of a *Thèse d'Etat*, the achievement of a lifetime and the result of a decade of research in France's original model of an academic career before the mid-eighties. As a whole, Jean Tournier's production is not limited to this scholarly thesis. Companion textbooks consisted of a simplified textbook (Tournier 1988, now republished by ellipses, Tournier 2005) and a lexicon (Tournier 1987). The 2007 reprint was both a long-awaited and an important contribution to the field of English word-formation in France for the last decades. This is all the more notable as lexicology in English departments is a *parent pauvre* as evidenced by the relative scarcity of lexical morphologists of English in France (Michel Paillard in Poitiers, Christian Bassac and Pierre Arnaud in Lyon have retired in the last five years). Yet the 'as is' or 'as was' reprint of the 1985 edition leaves room if not for deception at least for updating. Even

then, in spite of now striking absences (statistical approaches to productivity, Prosodic morphology to name but a few), the book remains descriptively important and interesting. Up to a point, his contribution to the general architecture of word-formation processes has some common points with the enunciative tradition, the long dominant framework in English departments in France (see Ballier 1997, Groussier 2000 or Valette 2006 for an overview).

The representation of the lexicon enacted by the lexicogenic theory of Tournier can be summed up by the corpora (dictionaries analysed) and the matrices presented. As the book cover indicates, much is said about the methodological unity and the desire to offer a descriptive approach of the word-formation processes within a conception of lexicogenesis. It can be deemed to be a typological approach of word-formation. His synthetic chart on page 51 is worth reproducing here as it recaps word creation / coinage (*néologie*) as analysable in twelve “matrices”

Table 1 The morphological matrices (after Tournier 1985)

Morpho-semantic coinage	Construction : affixation Compounding Phonic motivation	Prefixation Suffixation Backformation Juxtaposition Blends Onomatopoeia	Internal matrices
Semantic coinage	Class-transfer (class shift) Metasemy (meaning shift)	Conversion Metaphor Metonymy	
Morphological coinage	Signifier reduction	Aphaeresis Apocope Initialism	

Borrowing	Morpho-semantic borrowing Semantic borrowing and calque Morphological borrowing	External matrix
-----------	---	--------------------

A terminological point is in order. On a par with the *phoneme / allophone* distinction, *lexemes* are the abstract virtual units, *lexies* are the contextualised tokens. In his book, Tournier has *lexies* rather than *lexemes*, which I have rendered as “lexical units”. The interdependence of the phenomena (matrices) is described, as well as the possible succession of the morphological processes.

With some respect, and with all due respect, this work is very much a pre-corpora book and pre-internet investigation, yet some 75, 000 lexemes (*lexies*) were taken into consideration. In the index, synapsies is mentioned in the analytical bibliography (461) but not quite exploited for the study of English words. (Tournier 1985:175-6) analyses conglomerates as form of “vertical conversion”, in an intellectual move influenced by Tesnière, a lexicalized transfer from a structural level to another (sentence, clause, group, morpheme) and distinguishes upgrading (affixes becoming words such as an anti) and downgrading (sentences downgraded to words, *a hug-me-tight*). By current standards, the book somehow overlooks IT facilities (CD-Roms) and corpora. The underestimation of collocation is sensitive in or the use of expressions such as the “surroundings of the words” (“l’entourage du mot”, discussing Galisson 1979). Sadly enough, though, the thesis was not updated for its reprint in 2007. It is not so much that the analysis has gone out of fashion, but examples lack luster when compared with recent coinages. Worst, the trends in word-formation exemplified have missed the internet turn and the renewed importance of reanalysis based on spelling it has fostered as I see it. I will outline some potential new venue for research by focusing on types of examples absent from Tournier as evidenced by the following examples.

Among recently published books are Kerry Maxwell’s *Brave New Words, A language lover’s Guide to the 21st Century* (Macmillan, 2006). This unassuming little book aimed a wide audience offers 200 examples, personal definitions and pronunciation of recent coinages and provide a set of recent and attested examples of new lexemes. Among the emerging patterns and phenomena that are not quite

accounted for by Tournier's framework are the ever-growing increase of spelling and the rise of submorphemic morphs arising from the pronunciation such as *a dotcom* (which shows the back and forth motion from sounds to letters and back, since “dot” literally spells out the graphic sign. Some of the alluring examples from Maxwell 2006 are *freegan* (coined after vegan), *blooks* (potentially a blend from *blog* and *book* to designate these blogs turned into books. As to the emergence of sub-acronym morph, *IM* for *Instant Messenger* could be a good candidate.

It is not surprising that IT might call for a necessary re-assessment of the role played by spelling in word formation, probably beyond Crystal's *Texting the great deb8*. In Saussurean terms, this allows the substitution of signifiers on phonological grounds (eight /eit/) but that substitution entails a change in the signified. Initial substitution of the first letter offers new paradigms for blends, see for instance *Hollywood* / *Bollywood* and *Nollywood*, and even *Tollywood* and *Lollywood*. (*Nollywood* is Nigeria's *Bollywood*. *Bollywood* is *Bombay's Hollywood*). While *Bollywood* retains some of the classic features of blends, with the common Onset and nucleus of the first syllable, *Nollywood* is almost nothing but a reanalysis of the initial consonant. Another case in point would be *pantyhose* / *mantyhose* — *mantyhose* is the male counterpart of *pantyhose*. While *pantyhose* is said to be formed as a compound (*panty*/), such truncation as *pants* and *panties*, the existence of *pants* testify to the relevance of reanalysis. It is my claim that the Internet has triggered the multiplication of such formations, playing with the limits of phonological constraints on a par with the desire to create IT-connotated terms. Blends is a playing field of the utmost interest, where creations such as *vlog* / *blog* are not only blends on the initial syllable, but possibly paradigmatic substitutions of letters entailing violations of English phonotactics. As a phonologist by trade, I am constantly struck by the interplay between spelling and reading. An instance of this is provided with W/w, a *dub-dub*, the clipped form of “doubleyou”, for W/W waiter / waitress. Another consequence of the importance of the graphic code is the blurring of the boundary between acronyms and initialisms. Word-formations like VPILF (SARAH PALIN - VPILF A Vice President I'd like to...), or *gmilf* — an acronym for Grandma I'd Like To F..., an undeniable way of flouting English phonotactic rules, whereas the original *milf* is a readable sequence

consistent with English phonotactics (see *milk* and *pilfering*). There are paths to explore in the discourse / graphic production continuum, as well as the discourse/lexicon interface.

Benveniste's *Problèmes de linguistique générale* (1966:30) is quoted as a conclusion in Tournier's work but the exploration of the discourse/lexicon interface seems somewhat under-researched in this perspective. I would like to suggest that a delocutive matrix could be taken into consideration, were we to adopt this matrix conception.

6.2 *The delocutive matrix?*

This last subsection examines conglomerates in light of delocutive productions. They are a way of exploring the language/*parole* continuum, the linguistic coining of clichés, proverbs, what Umberto Eco calls the Encyclopedia of the speaker ("The encyclopedia is the regulative hypothesis that allows both speakers to figure out the 'local' dictionary they need in order to ensure the good standing of their communicative interaction", Eco 1986: 80). The lexicon is also partly structured by these cultural fragments, admittedly mentioned by Tournier in the guise of (debatable and disputed) cultural textual lexies, propagated in discourse.

Benveniste has defined delocutivity with his study on what he calls 'delocutive verbs': "We shall call "delocutives" those verbs which we propose to establish as *derived from locutions* (emphasis in the original). He takes the example of *salutare, salutem dare* (*to greet*) which comes from *salus!*. "Such a verb is defined by its relation to the formulaic expression from which it is derived and will be called 'delocutive' " (238). Similarly, he gives for English *to hail* (*to shout 'hail'*), *to encore* "*to shout encore*" and for American English *to okay*, and even *to yes*. From that point of view, the etymology of *safari* (from Swahili *safar* (*travel*, OED) could be translated by 'have a good trip'. In other words, a speech act turned into a word. As Benveniste notes in his conclusion:

The essential and signal feature of a delocutive is that it is in the relationship of "to say..." with its nominal base, and not in the relationship of "to do...", which belongs to the denominative. It is not the least instructive characteristic of this class to show us a

sign of language deriving from a locution of discourse and not from another sign of language; by this very fact, delocutives are, above all, from the moment at which they are created, verbs denoting activities of discourse. Their structure as well as the reason that summon them into existence assign them a very particular position among the other classes of verbal derivation” (246)

One could argue that conversion suffices to explain *hail*, *yes* or *okay* but the point made about discourse and the dissemination and lexical units still holds. Within this delocutive matrix, other phenomena could be included such as catchphrases, format phrases and snowclones.

Conglomerates are part of a more general cline for delocutive word-formation spotted by Emile Benveniste in his seminal paper about new forms of (nominal) compounding where referents are referred to in the linear order? Would it be possible to posit a speech / discourse cline, *from discourse to language (citation/proverb/format phrase)?*

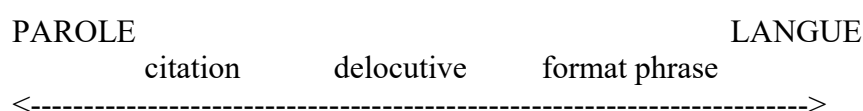


Figure 2 lexical units at the discourse / language interface

Arguments in favour of this delocutive matrix, or at least of this discourse / lexicology interface can be found in the examples given by the OED for the definitions of the conglomerates *oops-a-daisy* or *fellow-well-met* (“to be (hail) fellow well met : to be on terms of free and easy companionship with (a person”, OED) :

- (12) 1581 G. Pettie tr. S. Guazzo Ciuile Conuersat. (1586) iii. 171 Being as you say haile fellow well met with his servant.
- (13) 1858 T. P. Thompson Audi Alteram Partem I. xxxvi. 137 The High Church Tory...offers...to be fellow well met with any of them.

- (14) 1885 W. J. Fitzpatrick Life T. N. Burke I. 308 The best fellow-well-met in the world.

In the examples given by the OED, the sequence is conveniently for my argumentation first used in a citational context (12), a copular predicative construction (13) and as an NP (14), skipping predicative uses in-between.

Conglomerates may in turn become productive patterns within limited paradigms (see *tug-of-war* for tug-of-love), a point made for Italian by Michele Cortelazo 2000 : 202 (quoted in Radmynski 2006), *va e veni* (comes and goes) *mangia e bevi* (eats and drinks). Similarly, for examples of variation on a limited paradigm, the OED has *honest-to-goodness*, *honest-to-God*, *honest-to-Christ* or *follow-my-leader* and *follow-the-leader*. As Meibauer 2007 puts it about “phrasal compounds”, these phenomena “nicely illustrate the possibility of insertion from syntax into morphology.”

In that sense, conglomerates are not unlike format phrases or catchphrases, building blocks of the Encyclopedia of the speaker. As Eco (1986: 80) puts it “A natural language is a flexible system of signification conceived for producing texts, and texts are devices for blowing up or narcotizing pieces of encyclopedic information.” Literary allusions, quotations (*once more, to the breach*), proverbs, and what Nigel Rees calls “Format phrases” (Rees 1990), idiomatic patterns where a single slot only is likely to allow for variation, often after a film title as *the year of V-ing dangerously*. My favourite is *the winter of discontent* to refer to strikes after 1979, which has been extended to *summer of discontent*, showing the productivity of the micro-paradigm *season of discontent*. These are truly syntactic matrices, possibly based on film titles, quotes and cultural allusions, as with Richard III’s opening monologue.

7 Conclusion

This paper has argued that conglomerates as delocutive outputs ought to be taken into consideration in the analysis of the lexicon and that Benveniste's innovation should not be disregarded, especially because the term can be used to characterize languages so different as Bambara, Slovak, English and French. Although construction grammar has many assets to deal with this kind of linguistic phenomena, the conclusion would like to

make the case for the relevance of an utterance-based grammar, where these complex NPs are uttered / pronounced in context, delineating the intellectual background of French enunciative linguistics (Benveniste, Antoine Culioli and, up to a point, Gustave Guillaume) that fostered such a theoretical proposal as “conglomeration”. It seems to me that these morphological outputs exemplify the usefulness of “utterance grammar” (*la linguistique énonciative*) among usage-based models (see Barlow & Kemmer 2000) and these forms of compounding / multi-word units are tokens of the particular interest of the utterance as a unit fit for linguistic analysis. How do objects of discourse make it into the lexicon? Benveniste’s concept of conglomeration can be used as a cogent notion to explore this discourse/morphology interface (Gaudin & Guespin 2000, Fradin 2003, Gaeta & Ricca 2009). I sometimes rant that we have the research questions of our theoretical frameworks, I would readily contend that conglomerates pertain to the whole utterance grammar tradition.

It is high time linguists of English working in France accepted an aggiornamento of the terminology they have used when (however scarcely) writing in English about enunciative linguistics. First-generation linguists within this tradition have used source-oriented translations into English and have mostly met with raised eyebrows. Target-oriented translations should be the real order of the day, if necessary with provisos explaining why an English mainstream term fails to capture the flavour of the original theoretical stance. In this sense, I contend that “utterance grammar” might sound more appealing to an English-speaking audience than “enunciative linguistics”.

References

- Arnaud, Pierre. 2004. *Le nom composé, Données sur seize langues*. Lyon: Presses universitaires de Lyon.
- Baayen, Hans. R. & Piepenbrock, Richard & Van H, Rijn. 1995. The CELEX lexical database (release 2 ed.): Linguistic Data Consortium.

- Ballier, Nicolas. 1997. *Les écoles françaises de linguistique anglaise*. Unpublished PhD Thesis under the supervision of Jean-Jacques Lecercle. Nanterre: University Paris X-Nanterre.
- Ballier, Nicolas. 2004. *Praxis métalinguistiques et ontologie des catégories*. Unpublished Habilitation Thesis under the supervision of Jean-Jacques Lecercle, University Paris X-Nanterre.
- Barlow, Michael & Kemmer, Susanne. 2000. *Usage-based models of language*. Stanford: CSLI publications.
- Bassac, Christian. 2004. *Principes de morphologie anglaise*. Bordeaux : Presses Universitaires de Bordeaux.
- Benveniste, Émile. 1966. Différentes formes de la composition nominale en français. In *Bulletin de la Société de Linguistique de Paris*, vol. 61. 82–95.
- Benveniste, Émile. 1974. *Problèmes de linguistique générale*. Paris: Gallimard.
- Boucher, Paul. 1992. Teaching English in compounds in English: an application of the Charlie Brown principle. In *L'ordre des mots – domaine anglais Travaux LXXVI*, Saint-Etienne Université de Saint-Etienne. 33-50.
- Creissels, Denis. 2004. Bambara. In *Le nom composé, Données sur seize langues*. Lyon: Presses universitaires de Lyon. 21-46.
- Dubois, Jean & Dubois, Claude. 1971. *Introduction a la lexicographie : le dictionnaire*. Paris: Larousse.
- Dumestre, Gérard. 2003. *Grammaire fondamentale du bambara*. Paris: Karthala Editions.
- Eco, Umberto. 1986. *Semiotics and the philosophy of language*. Bloomington: Indiana University Press.

- Fiala, Pierre. 1989. Figements et phraséologie. État des recherches actuelles. In *Courants sociolinguistiques*, 8. 137-157.
- Fradin, Bernard. 2003. *Nouvelles approches en morphologie*. Paris: PUF.
- Gaeta, Livio & Ricca, Davide. 2009. Composita solvantur: Compounds as lexical units or morphological objects. In *Compounds between syntax and lexicon. Special Issue of Italian Journal of Linguistics / Rivista di Linguistica*, 21(1). 35-70.
- Groussier, Marie-Line. 2000. On Antoine Culioli's theory of enunciative operations. *Lingua* 110.3. 157-182.
- Halaoui, Nazam. 1993. Du sens au signe: La néologie des langues africaines en contexte terminologique. In *Meta: journal des traducteurs / Meta: Translators' Journal*, vol. 38, n° 4. 704-718.
- Hučka, Pavol. (in prep). Conglomeration in English - a corpus-based investigation. Dissertation thesis jointly supervised by Prof. Štekauer and author.
- Mahrer, Rudolf. 2011. Vers une linguistique de la parole, à partir de Benveniste. In Brunet, Emilie & Mahrer, Rudolf (eds). *Relire Benveniste: Réceptions actuelles des Problèmes de linguistique générale*, 3. Louvain: L'Harmattan. 197-210.
- Labbé, Dominique. 1990. Normes de saisie et de dépouillement des textes politiques. In *Cahier du CERAT*, (7). 1-135.
- Léon, Jacqueline. 2004. Lexies, synapsies, synthèmes: le renouveau des études lexicales en France au début des années 1960. *History of Linguistics in Texts and Concepts. Geschichte der Sprachwissenschaft in Texten und Konzeptionen*. Gerda Hassler (ed.) Münster: Nodus Publikationen. 405-418.
- Masini, Francesca. 2009. Phrasal lexemes, compounds and phrases: A constructionist perspective. *Word Structure* 2 (2). 254-272.

- Meibauer, Jörg. 2007. How marginal are phrasal compounds? Generalized insertion, expressivity, and I/Q-interaction. In *Morphology*. 233-259.
- Nosek, Jiří. 1985. Quotational Compounds in Modern English and Czech. In *Philologica Pragensia* 28. 159–164.
- Paillard, Michel. 2000. *Lexicologie contrastive anglais-français: formation des mots et construction du sens*. Paris: Ophrys.
- Ryšavá, Andrea. 2012. *Noun phrase in English: its form, function and distribution in text*. MA Thesis [Vedoucí práce: Prof. PhDr. Jarmila Tárníková, CSc.] Oloumouc.
- Radimský, Jan. 2002. *Les composés italiens actuels*, Paris: PUF.
- Rees, Nigel. 1990. *Bloomsbury dictionary of popular phrases*. London: Bloomsbury.
- Rendell, Ruth. 1994. *Simisola*. New York: Arrow.
- Reichs, Kathy. 2007. *Bones to ashes*. New York: Arrow.
- Schwarze, Christoph. 2006. Processus grammaticaux dans la formation des mots. Judith Meinschaefer & Katrin Mutz (eds). Arbeitspapier Nr. 120 Morphologie und romanistische Sprachwissenschaft Carmen Kelling, Morphologie und romanistische Sprachwissenschaft. Akten der gleichnamigen Sektion beim XXIX. Deutschen Romanistentag: Saarbrücken. 239-261.
- Sebeok, Thomas Albert (ed.). 1972. *Current trends in linguistics: linguistics in Western Europe*. The Hague: Mouton.
- Tournier, Jean. 1985. *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Paris: Champion.

- Trips, Carola. 2012. Empirical and theoretical aspects of phrasal compounds: against the "syntax explains it all" attitude. In Ralli, Andrea (ed.). *Proceedings of the Mediterranean Morphology Meeting 8*. 322-346.
- Trips, Carola. 2014. How to account for the expressive nature of phrasal compounds in a conceptual-semantic framework. *Skase Journal of Theoretical Linguistics*, 11:1. 33-61.
- Valette, Mathieu. 2006. *Linguistiques énonciatives et cognitives françaises (Gustave Guillaume, Bernard Pottier, Maurice Toussaint, Antoine Culioli)*. Paris: Bibliothèque de grammaire et linguistique.
- Villoing, Florence. 2012. French compounds. *Probus*, Volume 24, Issue 1, Berlin: De Gruyter. 29-60.
- Wells, John. 1990. *Longman Pronouncing Dictionary*. London: Longman.

Nicolas Ballier
Université Paris Diderot (Sorbonne Paris Cité)
CLILLAC-ARP (EA 3967)
UFR d'études anglophones
Bât. Olympe de Gouges, Case 7046
5 rue Thomas Mann
75205 Paris Cedex 13
e-mail: nicolas.ballier@univ-paris-diderot.fr

**The lexical/phrasal status of Polish
Noun+Adjective or Noun+Noun combinations
and the relevance of coordination as a diagnostic test**

Bożena Cetnarowska, University of Silesia

Abstract

*This paper considers the morphosyntactic status of Noun+Adjective and Noun+Noun combinations in Polish, such as *dział finansowy* (lit. department financial) ‘a finance department’ and *pilot-oblatywacz* ‘a pilot-navigator’. A brief review is offered of diagnostic tests which can be employed to determine the phrasal or lexical status of N+A and N+N units in Polish. Reference will be made to various criteria of compound-hood proposed for English (e.g. by Bauer 1998, Giegerich 2005 and Lieber and Štekauer 2009). Particular attention will be given to the predictions of the coordination test.*

Keywords: *juxtapositions, phrasal lexemes, coordinated compounds*

1 Introduction

Speakers of Polish frequently put classifying adjectives after head nouns, to form such multiword expressions as those given in (1-3). The adjectives in such N+A combinations are often (though not exclusively) denominal relational adjectives, e.g. the lexeme *finansowy* ‘financial’ is derived from the noun *finanse* ‘finances’.

- | | | |
|-----|------------------------|-------------|
| (1) | dział | finansowy |
| | department | financial |
| | ‘a finance department’ | |
| (2) | ogród | zoologiczny |
| | garden | zoological |
| | ‘a zoological garden’ | |