

Comparative Overview on Word-Compounding in English and Albanian

M.A. Esmeralda Sherko

Faculty of Veterinary Medicine
Department of Morpho-Functional Subjects, ESP Division
Agricultural University of Tirana
e-mail: aldasotiri@yahoo.com

Abstract This paper presents a comparative bilingual overview, by comparing word formation in two languages, respectively Albanian and English, as well as the main word formation trends in both languages. Both languages display a variety of word formation techniques, but their classifications according to different ways are not in complete conformity with each other. An example of this is "përngjitja", which in Albanian is classified as an independent class form the compound nouns, while in English it appears under the heading of "compounds". The paper will also highlight spelling and phonetic differences in both languages. In Albanian language the "përngjitja" is formed by the monomial of word group or phrase, a prepositional phrase or a group of words in general. The syntactic relations between the terms of the phrase or the group are usually preserved clearly, even in the structure of the newly formed word -të përngjitur. In English compounding, which also includes përngjitja refers to the process of combining two words to form a new one, where the words added at the beginning to a certain extent subcategorizes the words following it.

Key words: përngjitja, word-formation, Albanian, English, language

1. Introduction to word formation

Word formation is the branch of linguistics that examines derived words synchronically and diachronically and in terms of their origin and functioning, their productivity or unproductivity, as well as their use in different styles of speech.

This study focuses on the main word formation processes in English and Albanian. While both languages possess a wide range of word formation techniques, careful inspection reveals that not all of them coincide.

► Word formation in Albanian²

The most common word formation processes in Albanian language are: derivation and its subdivisions (prefixed derivation, suffixal derivation, prefixed-suffixal derivation, derivation without apposition), articulation, composition, përngjitja, conversion, and the mixed manners:

- a) përngjitje and prefixation
- b) composition and suffixation
- c) articulation and suffixation
- d) articulation and prefixation- suffixation
- e) articulation and përngjitje

The above mentioned word formation techniques are divided as follows:

- a) morphological ways, which include: affixation, articulation, composition
- b) non-morphological ways: përngjitja and conversion.

► Word formation in English³

The most common word formation processes in the English language are: derivation, agglutination, back-formation, blending, acronyms, clipping, calque, semantic loan, compounding, incorporation, conversion, neologism, loanword,

¹ This term does not have a corresponding equivalent term in English language.

² Grammar of the Albanian language I, Tiranë 1995

³S.; Quirk. R. "A student's grammar of the English language".

onomatopoeia.

2. Overview on word –compounding in Albanian

One of the most prolific ways of word formation is derivation and its subdivisions as well as composition, which are closely interrelated (E. Hysa⁴). A compound word consists of two or more motivating themes, which are used independently even without this composition, as two meaningful separate parts of speech.

Another important word formation technique in Albanian language is called – “Përngjitje”⁵. It is used to define new words formed by the monomial of a word group or phrase, of a prepositional phrase or a group of words in general. The syntactic relations between the terms of the phrase or the group are usually preserved clearly, even in the structure of the newly formed word -të përngjitur.

The formant of this word formation technique is the definite order of the components of the word (in words of related subordination, the subordinate component is generally placed at the beginning) and the common stress (whereas the other components are unstressed or have a secondary stress).

Most of the words “të përngjitura” derive from a monomial set phrase, or a prepositional phrase (preposition + noun). In the long continuous process of their use the terms of a word group are attached to each other, thus forming a unique word, where the morphological indices which have lost their grammatical functions might be preserved in a fossilized way. Considering the fact that the components these words no longer have their character of unique independent words, the newly formed word is new, not only by the lexical point of view, but also by the morphological – phonetic one. Such words- të përngjitura – are found in a lot of invariable parts of speech.

• The main characteristics of the these words (respectively –të përngjitura) when compared to compounding are the following:

- a) These words- (respectively –të përngjitura) – are always made up of specific words, that generally preserve clear syntactic relations between them;
- b) The meaning of the this word (respectively- e përngjitur), as a rule, is equal to the meaning of the group from where it came out;
- c) This word (respectively- e përngjitur) might have in itself even a single designating word, or might be made up mainly of words belonging to secondary parts of speech;
- d) These words (respectively- të përngjitura) are usually the result of merging a phrase or a group of words during a long process in use.

Despite these differences compounding and “përngjitja” are not unrelated or without relation between them. Historically composition derived from përngjitje. This transition is present even in today’s language (contemporary language) in the case of “përngjitje” of word groups-phrases of *noun + noun in ablative case*. In this case the final “i” letter characteristic of the ablative, no longer preserves its grammatical meaning as an ending, e.g. (“vajgursjellës” not “vajgurisjellës”). This would turn these nouns into authentic compositions.

Word formation techniques often intertwine with each – other. The most common ones are:

- a) **Përngjitje with prefixation.** This way forms *adverbs*, such as: *njëherazi*, *paprtimas* etc.; *nouns*: *nëpunës*, *mëditje* etc.; thus they are formed from the përngjitje of a prepositional group or an analytic verbal form, adding a prefix.

Other similar formations include adjectives formed by the merging of a *numeral group* (or the indefinite pronoun “shumë”) + *noun with the prefix–sh* (e), deriving from the inflection of the plural indefinite ablative case, e.g.: *pesëlekësh*, *gjashtërokësh*, *shumëkëmbësh* etc.

- b) **Composition with prefixation.** They are formed by the composition of a phraseological unit, to which a suffix is added, e.g.: *pjesëmarrës* (*marr pjesë*), *marrëveshje* (*marrë vesh*) etc., or adverbs, such as *këmbadoras* etc.

⁴ E. Hysa “Philological studies” 1986 – Affixes and their position in word formation Alb. language. Pg 91

⁵ Grammar of the Albanian language I, Tiranë 1995

This must be clearly distinguished by the authentic compositions. In some cases such as *liridashës*, *paqedashës*, this distinction is clear, because in Albanian there are no such words as “*dashës*, *veshje*” (formed by the noun “*vesh*”). The second terms of these compositions structurally resemble to the suffixal formations, but they have been formed within the rules of compositions, they do not really exist when separated from it, only potentially.

- c) **Articulation and përngjitje**. This group includes adjectives from the articulation of a prepositional phrase: *i pashkollë*, *i pashtëpi* etc.
- d) **Articulation and përngjitje and prefixation**. Adjectives formed according to this manner include a prepositional group (with the preposition *për*) and the suffix *-shëm*: *i përdyjavshëm*, *i përmuajshëm* etc.

• Përngjitja in different grammatical categories⁶

► **Nouns (të përngjitur)**. This category of nouns is limited. Such nouns are, e.g.: *farefis*, *gjëgjëzë*, *mirëseardhje*, *ecejake*, *thashetheme* etc.

The most common cases occur when a group with a preposition + noun is added to a prefix (përngjitje + prefixation), e.g.: *pagjumësi* (*pa gjumë + si*) *nëndetëse* (*nën det + se*) etc.

► **Adjectives (të përngjitur)**. As far as adjectives are concerned përngjitja is not a productive way. Such adjectives are: *bardhezi*, *kuqezi* etc.

► **Numerals (të përngjitur)** from *njëmbëdhjetë* up to *nëntëmbëdhjetë*. These words (*të përngjitura*) are different from the phrases, from the order of the compounding terms as well as from the connective word of the terms.

► **Indefinite pronouns (të përngjitur)**. When considering the indefinite pronouns of this category we distinguish that one of the elements of the indefinite pronoun- *të përngjitur* – is usually declinable, whereas the other is rigid: *dikush* (*dikujt*, *dikë*). But there are also indefinite pronouns of this kind where both elements are rigid: *çdo* (*ç-do*), *çfarëdo* (*çfarë-do*), *dicka* (*di- çka*) etc.

► **Adverbs (të përngjitura)**. A considerable part of adverbs is formed by - përngjitje. The largest group contains adverbs formed by the - përngjitja- of a word group, where one of the constituent parts is a pronoun. According to the elements there are several structural groups of such adverbs (thus *të përngjitura*).

- a) **Pronominal adverbs (të përngjitura)**, where the pronoun “*gjithë* “ is the first element of the - përngjitje, whereas the second element might be a pronominal adverb, a noun or another part of speech. E.g.: *gjithandej*, *gjithashtu*, *gjithmonë*, *gjithnjë*.
- b) **Pronominal adverbs (të përngjitura)**, where the first element of the - përngjitje – is one of the interrogative pronominal adverbs: *kur*, *ku nga*, *si*, *qysh*, *tek*, whereas the second element is the verb *do*. E.g.: *kudo*, *ngado*, *tekdo*, *sido*, *sado*.
- c) **Adverbs (të përngjitura)**, where the second element of the përngjitje – is the pronominal adverb *ku* in the form *kund* or *kundi*, whereas the first element might be a pronoun, an adverb, or a particle, such as: *asgjëkundi*, *askund*, *askurrkund*, *kurrkund*, *gjëkundi*, *tjetërkund*.
- d) **Adverbs (të përngjitura)**, where the second element of the compounding- përngjitjes – is the noun *herë*. The first element might be a preposition, a pronoun, an adverb, or a particle. E.g.: *asnjëherë*, *atëherë*, *menjëherë*, *ngaherë*, *nganjëherë*, *përherë*, *përnjëherë*.
- e) **Adverbs (të përngjitura) formed by a preposition and a noun**. This group includes: *përballë*, *përbri*, *përdhe*, *përfundi*, *përkrah*, *përmes*, *përçark*, *përreth*, *përditë*, *përmatë*, *përherë* etc. Obviously the second element of - përngjitje-, the noun is always in the indefinite form.
- f) **Adverbs (të përngjitura)**, whose constituent elements are two nouns merged through a vowel or not; or another part of speech might be a constituent element, more often an adverb. E.g.: *fytafyt*, *anekënd*, *turravrap*, *kryekëput*, *kryekreje*, *motmot*, *shpeshherë*, *rallëherë*, *rallëkund*, *buzagaz* etc. These formations might nowadays be considered compound words.
- g) **Adverbs (të përngjitura) formed by e numeral and the noun *fije* (*fill*) in indefinite ablative plural form, with the inflection *-sh*, such as: *dyfish*, *trefish* etc. As analogy to these stands even the word *njëfish*.**

⁶ Grammar of the Albanian Language I, Tiranë 1995

h) Adverbs (të përngjitura) having as constituent elements variable parts of speech. These include: *ndërkaq, sakaq, dosido, njëlloj, kësisoj, kësodore, sidokudo, vetvetiu* etc.

► **Prepositions (të përngjitura)**. These prepositions- përngjitura- are formed by the merging of two prepositions, (përngjitja – of two separate words) e.g.: *përmbi, nëpër, përveç*, merging the particle *si* with the prepositions *pas, mbas: sipas, simbas*, as well as those formed by conversion of adverbs - të përngjitura, such as: *përballë, matanë* etc.

► **Conjunctions formed by the –përngjitje- of two individual words**. According to the type of words participating in their composition, these conjunctions -të përngjitura –might be further classified as follows:

- a) The second element is a conjunction (generally one of the simple conjunction *sa, se, si, që* etc.), whereas the first might be:
 - also a conjunction: *po(r)sa, sesa, teksa* etc.
 - preposition: *mbasi (mbas +si), meqë, ngaqë* etc.
 - particle: *derisa, gjersa*.
 - adverb: *mirëpo*
 - verbal form: *meqënëse*
- b) the first element is a conjunction, whereas the second might be:
 - pronoun: *siç (si +ç)*
 - verbal form: *ndo (në +do)*
- c) none of the constituent elements is a conjunction; such are the conjunctions *megjithatë, megjithëkëtë*, formed by the – përngjitje- of the preposition *me* and the indefinite pronoun *gjithë* with the demonstrative pronouns *atë* and *këtë*, as well as the conjunction *gjithsaherë*, formed by an indefinite pronoun, an interrogative pronouns and a noun.

► **Particles (të përngjitura)**. The words: *pale, posi, (po)thujse* etc. They are limited in number. They are formed by the monomial of two or more words because of the long continuous use next to each-other. Among these words - të përngjitura – usually one is a particle or a conjunction. Thus there are:

- a) particle + verb → *pale = pa +le*
- b) particle + adverb → *posi = po + si*
- c) conjunction + pronoun → *seç = se+ ç*
- d) verb + conjunction → *thujse = thuj + se*

In rare cases there might be the monomial - përngjitje – of a whole sentence, such as the particle *kushedi*.

► **Interjections (të përngjitura)**. Simple interjections, especially those formed by a single vocal as well as those formed by two or more vocals, are often use repeatedly. This repetition has derived into their - përngjitje, or by the formation of several interjectory expressions, by thus enriching the fund of the Albanian interjections with new formations. Some of the non-simple interjections have been formed by monomial words, by compounding terms of a sentence, full or absent, because of the long continuous use as stereotypic formulas of greeting or congratulating, such as: *faleminderit, lamtumirë, mirupafshim, tungjatjeta* etc, which have an overall meaning and the value of an inseparable unit unable to be analyzed in individual terms. In these compositions the words have completely lost their grammatical functions, and in some cases they have also undergone phonetic changes.

3. Overview on word –compounding in English

Compounding⁷ consists in the combination of two or more (usually free) roots to form a new word. Compounding is a very common process in most languages of the world (especially among synthetic languages). The meanings of the

⁷ The Oxford handbook of Compounding, Rochelle Lieber, Pavol Stekauer

words interrelate in such a way that a new meaning comes out which is very different from the meanings of the words in isolation.

A common semantic classification of compounds yields four types:

- endocentric
- exocentric (also bahuvrihi)
- copulative (also dvandva)
- appositional

An **endocentric compound** (A+B denotes a special kind of B) consists of a head, i.e. the categorical part that contains the basic meaning of the whole compound, and modifiers, which restrict this meaning. For example, the English compound *doghouse*, where *house* is the head and *dog* is the modifier, is understood as a house intended for a dog. Endocentric compounds tend to be of the same **part of speech** (word class) as their head. (Such compounds were called *tatpuruṣa* in the Sanskrit tradition.)

Exocentric compounds (A+B denotes a special kind of an unexpressed semantic head), (called a *bahuvrihi* compound in the Sanskrit tradition) are hyponyms of some unexpressed semantic head (e.g. a person, a plant, an animal...), and their meaning often cannot be transparently guessed from its constituent parts. For example, the English compound *white-collar* is neither a kind of collar nor a white thing. In an exocentric compound, the word class is determined lexically, disregarding the class of the constituents. For example, a *must-have* is not a verb but a noun. The meaning of this type of compound can be explained as "(one) whose B is A", where B is the second element of the compound and A the first. A bahuvrihi compound is one whose nature is expressed by neither of the words: thus a *white-collar* person is neither white nor a collar (the collar's colour is a metaphor for socioeconomic status).

Copulative compounds (A+B denote the sum of what A and B denote) are compounds which have two semantic heads. E.g. *bittersweet*, *sleepwalk* etc

Appositional compounds (A+B provide different descriptions for the same referent) refer to lexemes that have two (contrary) attributes which classify the compound. E.g. *maidservant*, *actor-director* etc.

In **English**⁸, compound words have the following characteristics:

1. Compound words behave grammatically and semantically as single words.
2. Since compound words behave as units, between their component elements no affixes (whether inflections or derivations) can usually occur; inflectional suffixes can appear only after compound words. For example, *bathrooms*, *school*, *buses*, *water resistant*. **Exceptions:** *passersby*, *brothers-in-law*, *courts-martial*.
3. Compound words can be written in three different ways:
 - a) **Open**, i.e., with a space between the parts of the compound; e.g., *toy store*, *diving board*.
 - b) **Hyphenated**, i.e., with a hyphen (-) separating the elements of the compound; e.g., *flower-pot*, *air-brake*, *she-pony*.
 - c) **Solid**, e.g., without a space or hyphen between the component elements of the compound; e.g., *flowerpot*, *washrooms*, *pickpocket*.

Preference for a particular form of writing the compound word depends largely on lexicographical conventions and the variety of English use. For instance, hyphenation (i.e., separating the elements of a compound with a hyphen) is more common in British English than in American English. In American English, the tendency is to write the compounds open or solid (Quirk et al., 1985).

4. Compound words usually have the primary stress on the first element of the compound; e.g., "air-Æcrafts", "chewing-Ægum". This fact differentiates compounds from phrases that have the same elements and order as compounds. Phrases usually have their primary accent on the second (or nominal) element; e.g., a "red Æ coat vs. a "Red Æcoat.

⁸ A Comprehensive Grammar of the English Language. Longman, Quirk et al., 1985

5. The second element (or head word) of the compound usually determines the grammatical category to which the whole compound belongs. Following are a few possible combinations:

n + n = n;	sunrise, dancing girl, hand-shake,
v + n = n;	rattlesnake, call-girl, dance-hall.
adj. + n = n,	darkroom, eyebrow.
n + adj. = adj.;	airsick, bottle-green.
pron. + n = n;	she-pony, he-goat.
prep. + v = v;	overtake, undergo.
prep. + n = n;	onlooker, off-day.
adj. + adj. = adj.;	gray-green, Swedish-American.

However, there are some cases in which the headword does not determine the grammatical class of the compound; for example:

n + v = adj.;	e.g., man-eating, ocean-going, heartfelt.
adj./adv. + v = adj.;	e.g., hard-working, good-looking, dry-cleaned.
n + prep. = n;	e.g., passer-by, hanger-on.
v + (adv.) prep. = n;	e.g., show-off, holdup.
v + adv. = n;	e.g., have-not, get-together.

It is important to point out that some compound words are made up of a bound root (or 'special' combining form, (as Quirk et al. (1985) call it), e.g., socio-, psycho-, and a free root; e.g., socioeconomic, psychoanalysis, biotechnology. The compound may also consist of two bound roots; e.g., nephrolithotomy, pornography.

6. **Compounding is a recursive process**⁹; i.e., one compound itself may become a constituent of a larger compound; e.g., lighthouse keeper, living-room furniture.

Compounds can be either native or borrowed. Native English roots are typically free morphemes, so that means native compounds are made out of independent words that can occur by themselves. Examples *mailman* (composed of free root *mail* and free root *man*)

Compounds formed in English from borrowed Latin and Greek morphemes preserve this characteristic. Examples include *photograph*, *iatrogenic*, and many thousands of other classical words.

Some compounds have more than two component words. These are formed by successively combining words into compounds, e.g. *pick-up truck*, formed from pick-up and truck, where the first component, *pick-up* is itself a compound formed from *pick* and *up*. Other examples are *ice-cream cone*, *no-fault insurance* and even more complex compounds like *top-rack dishwasher safe*.

Rhyming compounds (subtype of compounds). These words are compounded from two rhyming words. Examples: *lovey-dovey*

There are words that are formally very similar to rhyming compounds, but are not quite compounds in English because the second element is not really a word--it is just a nonsense item added to a root word to form a rhyme. Examples: *higgledy-piggledy*

This formation process is associated in English with child talk (and talk addressed to children), technically called hypochoistic language. Examples: *bunnie-wunnie*, *Henny Penny*

Another word type that looks a bit like rhyming compounds comprises words that are formed of two elements that almost match, but differ in their vowels. Again, the second element is typically a nonsense form: *pitter-patter*, *zigzag*.

⁹ Handbook of general and applied linguistics, Zapata Becerra, A. A. (2000).

4. Conclusions

This paper mainly focuses on comparing word formation in both languages trying to identify the similarities and differences between both languages. While in Albanian the distinction between compounding and "përngjitje" is clearly cut, in English there is no distinction between these two word formation techniques; they belong to the same group. However Albanian language has compound nouns separately, also corresponding to English under the heading of compounds.

While compound words in English are classified into open, hyphenated or solid, which distinguishes their lexicography, in Albanian there are no similar rules; they are all written as one separate word. The concept of recursive compounds in English does not have a similar counterpart concept in Albanian, since if translated into Albanian they belong to another category, not at all related to word formation. While "përngjitja" appears in different grammatical categories where words undergo declination in several cases, a characteristic for Albanian language, in English words are joined together without being declined. Compound words in both English and Albanian are made up of words having motivating themes, meaning they could be used as separate words. This is not the same for the words "të përngjitura" in Albanian, which in English are part of compounds.

References

- E. Hysa- "Affixes and their position in word formation Alb. Language" -"Philological studies" 1986
Grammar of the Albanian Language I, Tiranë 1995
Quirk et al., 1985, A Comprehensive Grammar of the English Language- Longman
Rochelle Lieber, Pavol Stekauer, The Oxford handbook of Compounding
S.; Quirk. R. " A student's grammar of the English language".
Zapata Becerra, A. A. (2000) -Handbook of general and applied linguistics,