

Panagiota Kyriazi¹
Athena Research Center
Aggeliki Fotopoulou²
Athena Research Center

MULTIWORD EXPRESSIONS OF GREEK LANGUAGE: A CASE STUDY OF NON-REFERENTIAL CLITICS IN MWES

The aim of this paper is to present an endeavour of classification of clitics within verbal Multiword Expressions (MWEs) of Modern Greek that seem *prima facie* to be lacking reference. There have been two types of classifications; the first one is based upon the syntactic structure of MWEs according to Maurice Gross's Grammar-Lexicon theory, and the second one is structured according to the degree of clitics' analysability. Through a detailed description and analysis of a linguistic corpus of 132 verbal multiword expressions, a taxonomy of 2 main categories of clitics is suggested, i.e., non-analysable and analysable, followed by 5 subcategories in accordance with their degree of analysability resulting in the construction of the term "spectrum of analysability".

Keywords: MWEs, spectrum of analysability, clitics, degree of fixedness

1. INTRODUCTION

Multiword expressions (MWEs) constitute idiosyncratic word sequences that represent a single semantic unit with a specific phonological, morphological, and lexical structure comprised of at least two words (Calzolari et al. 2002; Sag et al. 2002; Baldwin & Kim 2010). Their main properties lie upon the *collation*, i.e., the noteworthy co-occurrence with which are found in the lexicon, the *discontinuity* among the components of the expression as there may be distance between them, the *non-compositionality* since the meaning of MWEs is holistically extracted, the level of *ambiguity* in semantic level in terms of the way each expression is

¹ p.kyriazi@athenarc.gr

² afotop@athenarc.gr

interpreted within a given context, and the *variability* which is related to the flexibility in the idioms' formulation (Constant et al. 2017).

The verbal multiword idioms are comprised of a verbal phrase (VP) and a clitic that seems to lack reference, yet they jointly create a semantically idiosyncratic expression that lies at the interface of grammar and lexicon. The clitics are particles empty of meaning but critical for meaning attribution; the absence of these particles leads to a complete lack of meaning or the compositional meaning of the expression's constituents. Therefore, clitics play an essential role in the creation of idiosyncratic meanings that often derive from youth slang as lexicalisations (Androutsopoulos 2000).

By virtue of the ostensible lack of clitics' reference, we proceeded to two different classifications of clitics: the first one constitutes a syntactic taxonomy according to MWEs' syntactic structure as per Maurice Gross' Lexicon – Grammar theory and the second one is based on the degree of clitics' analysability, resulting in the formulation of two main categories: non-analysable and analysable clitics.

The aim of this paper is to display a detailed analysis and description of the aforementioned categories in accordance with the syntactic and semantic properties of verbal MWEs with clitic in Modern Greek. The findings, emerging from the application of a series of lexical and morphological tests, appear to be confuting the above observation as not all clitics lack reference and they appear to have different functions.

1.1. Background

The theoretical framework being deployed for the classifications in question is that of Maurice Gross' model (1975, 1981, 1984), known as "Lexicon – Grammar" (LG). According to this model, a simple sentence comprised of basic structural elements (Subject – Verb – Object[s]) is the main syntactic unit with various distributional and transformational properties attached to it.

The methodological tool of LG allows the systematic categorisation of linguistic units according to syntactic and lexical criteria and the classification in homogeneous clusters in table format³. The main structural components of LG are sentences with free combinations (N0 V N1) e.g., έφαγα χυλοπίτες (έfaga chilópites) 'I ate pasta', fixed or idiomatic expressions (N0 V C1) e.g., έφαγα χυλόπιτα (έfaga chilópita) 'I blow sb out', and light/support verb constructions e.g., έφαγα χαστούκι (έfaga chastóyiki) 'I slap sb'. Therefore, this model can be better understood as a continuum of fixedness, on the one end being the free combinations and on the other the stereotyped expressions (Anastassiadis-Symeonidis, Fotopoulou & Kyriacopoulou 2020).

³ The symbols for the description of each category are listed in Appendix 1.

Previous research projects on MWEs (Anastasiadis-Symeonidis 1986; Fotopoulou 1993; Ralli & Stavrou 1998; Mini 2009; Fotopoulou & Giouli 2018; Markantonatou & Samaridi 2018) have shown the morphological, syntactic, and semantic characteristics of multiword expressions as well as the structural constraints and semantic impact on the meaning of an expression resulting from fixed or non-fixed entities. Specifically, there has been enlightening research on MWEs in Modern Greek and this can be verified by the broad range of terminology regarding this topic: *fixed phrases* (Fotopoulou 1993), *stereotyped expressions* (Anastasiadis-Symeonidis & Efthymiou 2006), *multiword lexemes/phrasal idioms* (Xydopoulos 2008).

1.2. Methodology

The data for the analysis comes from a corpus of 132 verbal MWEs detected in text corpora and resources from the undergraduate dissertation of Kyriazi (2020), the thesis dissertation of Fotopoulou (1993), the thesis dissertation of Mini (2009), the Dictionary of the Vernacular and Marginal Language of G. Katos, and the website Slang.gr that hosts lexical entries of colloquial language.

A large amount of the data constitutes idiomatic expressions and slang from everyday language coming from the sociolect of youth language (Androutsopoulos 2001), since conversational utterances and everyday expressions based on the mechanism of metaphor are mainly -and rapidly- developed within the networks of young people. By the same token, multiword expressions' essence lies within their figurative meaning due to the combination of specific lexical items, in a particular structure that differentiates from literal meaning.

Initially, the data was gathered and classified in table format according to their syntactic and structural properties. A thorough testing and analysis of expressions' features was conducted through a plethora of examples and a series of morphological and syntactic criteria. The outcome of the first methodological stage was the classification of MWEs in 7 syntactic tables in accordance with their structure.

Following, clitics within verbal MWEs were analysed through syntactic tests in order to define their degree of analysability and fixedness. During this stage, two main categories were composed: the non-analysable and analysable clitics, with 5 subcategories in total (2 in the former and 3 in the latter). Both classifications are structured in properties tables including horizontal lines with a lexical entry of a multiword expression and columns corresponding to the expression's properties⁴.

With respect to phrases' argument structure, the elements being checked are the acceptance or not of prepositional complements and whether a VP

⁴ A syntactic table of the present analysis is illustrated to present the methodology of the analysis in Appendix 2.

constitutes a light/support verb construction based on its grammatical and syntactic properties (Kyriacopoulou & Sfetsiou 2002; Fotopoulou et al. 2021) as well as the degree of replaceability of the verb in the expression with the verb *to be* (Moustaki 1995). Amongst the distributional features according to Harris' distributional grammar theory (1954, 1970), there is the existence of animate or inanimate subjects, the negation form, complements indicating body parts or adverbs, as well as transformational properties, such as symmetry and effacement.

1.3. Definitions

Before proceeding to the main analysis, it is important to clarify the definitions being deployed relating to MWEs. Starting with clitics, are defined as “inflectional pronoun items characterised by tonal dependence, albeit they appear to be syntactic units” (Philippaki-Warburton 1994: 53). To that end, weak pronoun forms cannot stand independently in speech, but they are bound to a phonological host, to which they attach and form a single syntactic unit (Ralli 2005; Mavrogiorgos 2010: 5).

Concerning clitics' function, they do not seem to behave as independent lexical units nor as affixes; instead, they entail dual citizenship displaying a hybrid behaviour depending on their degree of fixedness and analysability (Klavans 1982, cited in Spencer & Luis 2012: 38). Added to this, the phenomenon of clitic doubling is closely related to proverb pronoun's syntactic capabilities, as it describes “the structure in which a clitic and a noun phrase (NP) co-exist, with the former to be referring to the latter without considerable pause or distance in between” (Philippaki-Warburton et al. 2002: 175).

By analogy to the spectrum of fixedness, i.e., “a continuum extending from free sequences, collocations, and quasi phrases, to fixed expressions” (Anastassiadis-Symeonidis, Fotopoulou & Kyriacopoulou 2020: 11), the term “spectrum of clitics' analysability” is introduced based on the syntactic and semantic properties that pronouns display in combination with the VP. Therefore, clitics can be understood within a continuum of fixedness, being described within 5 distinct categories, demonstrating a varying degree of analysability.

2. MAIN ANALYSIS

2.1. Syntactic Classification

According to the first classification, 7 syntactic categories are structured including verbal MWEs with clitics. The syntactic structures emerging from the analysis are:

1. NO PPV V
2. NO PPV V C1

3. NO PPV V PREP (N1+C1) PREP N2=N(gen)
4. NO PPV V (C1+Adj) PREP N= N(gen)
5. NO PPV V C1 (E+PREP N1)
6. NO PPV V (E+C1) QuP
7. PPV V CO (E+N(gen))

Category (1) constitutes the basic syntactic structure including the subject, the proverb pronoun, and the verb.

E.g., *Την έβαψα αν μου ζητήσει το νοίκι σήμερα*

Tin évapsa an moy zitísi to noíki símera

'I painted it' (= I am screwed) if she asks for the rent today

Structure (2) differentiates from the above as there is the fixed complement of the verb, either an object or adverb, as illustrated in the example below with *λίμπα* (*limpa*) 'mess' and *ρημαδιό* (*rimadio*) 'ruins' being the fixed complements of the verb.

E.g., *Κάν' τε πέρα να περάσω, κάν'τε πέρα να διαβώ, μην τα κάνω όλα λίμπα, μην τα κάνω ρημαδιό*

Κάν'τε péra na peráso, κάν'τε péra na diavó, min ta káno óla límpa, min ta káno rimadió

Make way for me to pass, make way for me to cross, not to 'make everything a mess', not to 'ruin everything'

In structure (3), there are two prepositional complements of which one may take either a fixed or a non-fixed object (ex.1), and the other one can be inflected as a possessive personal pronoun (ex.2), being the indirect complement of the verb.

(1) *Κάθε φορά που εμφανίζονται δυσκολίες το βάζει στα πόδια*

Káthe forá pou emfanízonte dyskolίες to vázi sta pódia

Each time a difficulty comes up he 'runs off'

(2) *Μου την δίνει στα νεύρα κάθε φορά που παραπονιέται*

Moy tin díni sta névra káthe forá pou paraponiéte

He is 'getting on my nerves' every time he grumbles

Syntactic category (4) is comprised of a fixed object or adverb that may be omitted and a prepositional complement that can be inflected as a possessive personal pronoun being the indirect complement of the verb.

E.g., *Από τότε που ο Γιάννης την πρόσβαλε, του το κρατάει + αμανάτι/μανιάτικο*

Από τότε που ο Γιάννης tin prósvale, toy to kratái +amanáti/maniátiko

Since John insulted her, she 'holds it against him'

In structure (5), there is a fixed object that may or may not, be followed by a free prepositional complement.

E.g., *Δεν τα βγάζω πέρα* (*E + με τα παιδιά*), *με έχουν ζαλίσει*
Den ta bgázo péra (*E + me ta pediá*), *me éhoyn zalísi*
 I cannot ‘pull through’ (*E + with the kids*), they have made my head spin

Category (6) includes a fixed object that can be omitted and a subordinate clause that operates both as an explanatory attribute to the proverb pronoun and as a verb’s object.

E.g., *Μην το δένεις* (*E + σχοινί κορδόνι*) *ότι θα έρθω*
Min to dénis (*E + shini kordóni*) *óti tha érho*
 Don’t ‘tie it’ (*E + rope string*) that I will come (‘don’t take it for sure’)

Lastly, syntactic cluster (7) entails a fixed subject instead of an object; there may also be a possessive case of the subject, as shown in ex. 1:

E.g., *Δεν το βάζει ο νους του ανθρώπου* *για τι αλήτη μιλάμε*
Den to vázi o noys toy anthrópou *gia ti alíti miláme*
 ‘No brain of a human’ can understand what a scumbag he is

2.2. Classification of clitics’ analysability

According to Cacciari and Levorato (1998, 1999), semantic analysability, or semantic decomposability (Gibbs et al. 1989), refers to the degree of analysis of an expression depending on the semantic composability of its components, while taking into consideration the factor of recognising its figurative meaning based on the semantic identification of its structural constituents.

Similarly, in the present classification, we attempt to demonstrate the degree of clitics’ analysability within verbal MWEs through a series of syntactic and structural tests. The degree of clitics’ fixedness can be described in terms of the semantic criterion of a phrase’s compositional or non-compositional meaning, the lexical factor concerning a word’s paradigmatic variation and, lastly, the morpho-syntactic aspect relating to a particular form in which a component is accepted to attribute a specific meaning, that otherwise, could not be ascribed within a free sentence (Lamiroy 2003).

Therefore, two main categories of clitics in verbal MWEs are shaped through the analysis of the corpus: non-analysable and analysable. The weak forms of proverb pronouns (*το, την, το, τα*) [*to, tin, to, ta*] have a specific distribution within the verb phrase while creating a cluster of dynamic relationship with the verb (Roussou 2015).

2.2.1. Non-analysable clitics

The category of non-analysable clitics is comprised of subcategories 1 and 2 and its dominant characteristic is the ultimate fixedness that clitics display in

combination with the VP. The clitic creates a single syntactic unit with the VP without having a distinct syntactic role. Another essential feature is that clitics cannot be analysed to an equivalent NP; there is an absolute degree of fixedness that does not allow this kind of lexical or semantic analysability as at least one component is invariable. The meaning of MWEs is non-compositional since it is holistically extracted from the aggregate of the lexical constituents. In other words, the meaning is understood as a whole rather than being built up from the meanings of the individual parts of the expression.

Subcategory 1

The clitics included in subcategory 1 are *το, τον, την, τα* (*to, ton, tin, ta*) and they construct a fixed non-analysable unit in conjunction with the VP. Their semantic core is detected within this conjunction and not in the individual components. For example:

- (1) **τον** ήπια (*ton ípia*) ‘I drunk it’ (= I am screwed) → ***τα** ήπια (*ta ípia*) ‘I drunk them’

Through the lexical criterion of substitution, we observe that the replacement of *τον* with *τα* has a semantic effect on the meaning; *τα ήπια* (*ta ípia*) is accepted as a free sentence with literal meaning (i.e., I drunk a lot of alcohol), whereas the meaning of *τον ήπια* (*ton ípia*) is figurative (i.e., I am screwed).

- (2) ο εκτυπωτής **τα** πίνει (*o ektipotís ta píni*) ‘the printer drinks them’ (=it is not working properly) → ο εκτυπωτής (***E+τα**) πίνει (*o ektipotís *E+ta píni*) ‘the copy machine drinks’

In case we completely omit the proverb pronoun *τα* through the effacement criterion, there is no accepted meaning at all. Here, the clitic is not only invariant but also a prerequisite for meaning attribution.

- (3) **τα** φτύνω (*ta ftíno*) ‘I spit them out’ → μισό χιλιόμετρο κάναμε και **τα**’φτυσε ***τα** σάλια (*misó hiliómetro káname ke ta’ftise *ta sália*) ‘we did half kilometre and he spit them out *the saliva’

The clitic pronoun *τα* cannot be analysed into the NP (*σάλια*). The clitic creates a stereotyped expression with the VP that is semantically fixed and not analysable into lexical counterparts.

Subcategory 2

In this subcategory, MWEs consist of verb phrases including only the proverb pronoun *την* (*tin*). The clitic is non-analysable and it forms a single syntactic unit with the VP in a stereotyped expression. The distinctive feature of this subcategory is that MWEs with only clitic *την* occur, with the latter referring to a generalised situation. Specifically, the term *κατάσταση* (*katástasi*) ‘situation’ is the referent phrase of the clitic and there is a morpho-syntactic association of the grammatical

gender between the clitic (**την**) and the referent phrase (κατάσταση), i.e., the feminine gender.

It is important to note that the term ‘situation’ is merely the deictic centre (Bella 2015) of the clitic and not its analysable phrasal counterpart, as the word *situation* cannot be incorporated in the MWE; it is an implicated reference point for the clitic *την* as a situated context.

E.g., Τελικά πάλι καθαρή *την* έβγαλες **την* κατάσταση
 Teliká páli katharí *tin* évgales **tin* katástasi
 ‘After all you got away with her again’ *the situation

Likewise, we observe through the application of morpho-syntactic criterion the following:

(4) **την** έβαψα (*tin* évapsa) ‘I painted her’ (= I am screwed) → ***το** έβαψα
 (to évapsa) ‘I painted it’

By substituting the feminine pronoun *την* with the neutral *το*, the meaning of the expression is altered from figurative, i.e. I am screwed, to literal, i.e. I actually paint something. Therefore, the conveyed meaning is in interaction with the proverb pronoun since a change in the latter has a direct effect on the former.

2.2.2. Analysable clitics

The second main category consists of analysable clitics within MWEs with three subcategories to be structured through the implementation of morpho-syntactic and lexical criteria: subcategories 3, 4, and 5. The distinctive feature of this category is the ability of clitics to be analysed into an NP due to their argument structure and the relation with the rest of the expression’s components as a result of transformational procedures (Foufi 2012).

The degree of fixedness of the MWEs is much looser than in the previous category thanks to the clitic’s reference and analysis into an NP. We observe instances of clitics with syntactic roles in relation to the verb, e.g., being a VP’s complement, or operating as an antecedent/anticipatory pronoun. MWEs’ meaning is compositional as each lexical unit refers to a semantic entity outside the phrase and it can be analysed.

Subcategory 3

In this subclass, the clitics *το*, *την*, *τα* (*to*, *tin*, *ta*) can be further analysed into NP or entire sentences as parts of speech. The clitics are syntactically and semantically important for the VP as they are responsible for the verb’s argument structure operating as a glue for the constituents of the MWE.

(5) επειδή άργησα *το* πρωί να πάω στη δουλειά, με κάλεσε ο διευθυντής στο γραφείο του και **τ’άκουσα** (*epidí árgisa to proí na páo sti doylíá, me kálese o diefthintís sto grafío toy ke t’ákoysa*) ‘because I was late

this morning at work. my boss invited me into his office and **I heard them'** (= I was reprimanded)

In ex. (5) clitic *τα* operates as a cognate object with the sense of the internal object that is semantically and interpretatively related to the action being described by the verb (Roussou 2015: 125). Therefore, *τα* refers to a sequence of words and, specifically, rebukes, being directly connected to the verb's meaning (*ακούω* ~ hear).

Another operation that clitics of this subclass may accomplish is that of an antecedent or anticipatory pronoun. Specifically:

(6) *δεν το βλέπω να έρχομαι* (*den to vlépo na érhome*) 'I don't see it coming'

In (6), the clitic pronoun *το* functions as an anticipatory pronoun as it refers to something that is yet to be mentioned in the trajectory of the sentence. According to Tzartanos (1930), an antecedent or anticipatory pronoun is "a pronoun through which the concept of a person or an object recurs or is anticipated and for which a reference or a mention will be accomplished later in speech".

The anticipatory clitic *το* adumbrates the existence of a reference point -a demonstrative reference- to understand the anaphora of the pronoun proverb. Therefore, the underlined subordinate clause *να έρχομαι* operates as an explanatory determiner to the clitic being an inextricable component for reference clarification and attribution of the corresponding meaning. In cases like these, we are talking about interlocutory elements dependent on the NP or sentence to which they refer; the cluster of the VP along with the clitic can construct ostensibly independent responses within conversation as long as the clitic's reference has already been made.

Subcategory 4

The clitics of subclass 4 operate as substitutes for the prototypical phrasal complements of the VP that have been effaced. The grammatical gender of the proverb pronoun is in agreement with the effaced NP and in some instances, the phenomenon of clitic doubling occurs without being a prerequisite for the assignment of MWE's meaning.

(7) *την* έκατσα την βάρκα → e.g., *αν μάθουν ότι εσύ μίλησες, την έκατσε*
tin ékatsa tin várka e.g., *an máthoyn óti esý mílises, tin ékatses*

'I sat it the boat' e.g., if they find out that you talked, you're screwed

(8) *την* πληρώνω τη νύφη → e.g., *θα αφήσεις έναν αθώο άνθρωπο να την πληρώσει έτσι;*

tin pliróno ti nífi e.g., *tha affsis énan athóo ánthropo na tin plirósi étsi?*

'I pay her the bride' e.g., will you let an innocent man take the fall like this?

The prototypical complements *την βάρκα* and *τη νύφη* of the VPs have been effaced and replaced by the clitic pronoun in the corresponding grammatical gender due to the morpho-syntactic trace of the object being left behind in that position. The clitic doubling structure is possible in phrases like these, as we come up with instances of the clitic pronoun being present and also co-referring with the NP in the same MWE e.g., *την πληρώνω τη νύφη ~ την έκασα την βάρκα*.

- (9) **τα** κρατάω μέσα μου (ενν. τα συναισθήματα) e.g., μην τα κρατάς μέσα σου, πρέπει να ξεσπάσεις κάποια στιγμή
ta kratáo méssa moy (mean. ta sinesthímata) e.g., min ta kratás méssa soy, prépi na ksespásis kápia stigmí
 'I hold them inside me' (mean. the emotions) e.g., don't hold them inside you, you need to burst out eventually

In (9), the prototypical complement is implicated, and so, there is no possibility of clitic doubling as the NP cannot be co-present with the clitic within the MWE; it has been effaced leaving the clitic behind as a trace.

Subcategory 5

In the last subcategory, we discern a class of VPs including only the clitic *τα* that has general reference: it refers to a series of things, actions, measures, or activities without specific anaphora. The analysability of clitic *τα* lies upon the fact that it is analysed into the word *όλα*, being the prototypical complement of the VP that can be either omitted or not.

Again, we observe that clitic *τα* has a number of identical grammatical features with its referent *όλα*, with the differentiation that the clitic pronoun not only refers to the NP but is also analysed to it. For example:

- (10) **τα** φέρνω βόλτα με τις δουλειές του σπιτιού παρόλο που είναι δύσκολο με τρία παιδιά
ta férho nólta me tis doylíes toy spitióy parólo pou íne dískolo me tría pedía
 'I bring them walk with the house chores even though it is hard with three children' (= I cope with etc.)

The clitic binds the meaning of the expression due to its morphological and grammatical features referring to a series of undefined things, i.e. I generally cope with house chores. This indefinite footprint is because of the relation between the clitic and its implicated analysable counterpart *όλα* ('all') that reflects this generality.

- (11) **τα** βάφω (E + **όλα**) μαύρα
ta váfo (E + óla) mávra
 'I paint them (E + all) black' (= I am disappointed)

In (11), the grammatical and syntactic agreement between the clitic *τα* and the object of the VP *όλα* is apparent; the fixed complement *μαύρα* operates as a

predicate to the clitic and, by virtue of that, the clitic, its referent, and its predicate agree in case (accusative), gender (neutral), and number (plural).

3. DISCUSSION

Through the analysis we conducted and the series of morpho-syntactic and lexical tests we applied we came up with two types of classifications. The first one involves the formation of 7 structural categories in properties tables according to M. Gross’ LG model and the second illustrates the structural and semantic constraints created by the VP and the clitic while discovering several patterns and distinctive properties that clitics display within a phrasal expression.

Specifically, we categorised proverb pronouns into two groups: non-analysable and analysable. The main characteristic of the former is the semantic and syntactic fixedness of the clitic in combination with the VP, as dictated by lexical and morpho-syntactic limitations derived from their relationship. In addition, the clitics of this class cannot be further analysed to an equivalent NP being lexically and semantically fixed. With respect to semantics, the meaning of these phrasal expressions is non-compositional.

The latter category includes the analysable clitics that are analysed and refer to an NP, that can either be present or not within an MWE. They are distinguished by their syntactic autonomy as they can function as complements of the phrase or as antecedent/anticipatory pronouns within speech being conversational elements in speech. Their meaning is compositional with each lexical component to be carrying recognisable parts of the expression’s idiomatic meaning (Nunberg 1994).

Thus, 5 subcategories emerge in total from these two classifications depicted in Figure 1 which encompasses the essence of the term “spectrum of clitics’ analysability”. A continuum of analysability is structured by moving from the polar of fixedness with non-analysable clitics to the other one of conventional expressions with analysable clitics. The pioneering element of this classification lies in the connection between an MWE’s fixedness and the degree of clitic’s analysability.

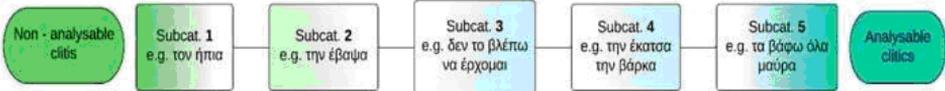


Figure 1.

4. CONCLUSION

The present paper aimed to give prominence to clitics' distinctive features and their several different categorisations according to their analysability in the context of MWEs; the clitics that *prima facie* seemed to be fixed without specific reference or phrasal counterparts to be analysed, appeared to have a degree of analysis within a phrase, a sentence, or a speech part.

Clitics' behaviour is described either as free in semi-fixed expressions or fixed in stereotyped expressions depending on the relationship built between the proverb pronoun and the VP's components. Therefore, the "spectrum of analysability" designates the degree of a clitic's analysability within a spectrum of fixed, semi-fixed and conventionalised expressions described by the 5 subcategories of non-analysable and analysable clitics we constructed.

Through this descriptive analysis of the fixed elements within speech, we structure a better understanding towards the semi-/non- fixed parts that together structure this spectrum of analysability. This kind of categorisation reveals important insights into the ways in which both structured written speech and conversational utterances are shaped and actualised in Discourse.

Acknowledgements

The present research project was conducted within the broader cooperation with the Institute of Language and Speech Processing (ILSP) and the Research Director Dr. Aggeliki Fotopoulou. In light of the general engagement with the syntactic classification and analysis of verbal multiword expressions, this project constitutes an attempt to categorise clitics of Modern Greek found in verbal MWEs based on their syntactic properties and their degree of analysability.

References

- Anastassiadis-Symeonidis & Efthymiou 2006: A. Anastassiadis-Symeonidis, A. Efthymiou, *Stereotyped expressions and teaching Greek as a second language*. Athens: Patakis Publication.
- Anastassiadis-Symeonidis, Fotopoulou, Kyriacopoulou 2020: A. Anastassiadis-Symeonidis, A. Fotopoulou, T. Kyriacopoulou, Multiword expressions in Modern Greek: synthetic review on their nature. In S. Markantonatou & A. Christofidou (eds.), *Bulletin of Scientific Terminology and Neologisms*, Special Issue: «MWEs in Greek and other languages: from theory to implementation».
- Anastassiadis-Symeonidis 1986: A. Anastassiadis-Symeonidis, *Neology in Modern Greek* [PhD Thesis] Thessaloniki: University of Thessaloniki. Scientific Yearbook of the School of Philosophy of Aristotle University of Thessaloniki, Annex no 65.
- Androutsopoulos 2000: J. Androutsopoulos, Extending the concept of the (socio) linguistic variable to slang. *Mi a szleng*, 109–140.
- Androutsopoulos 2001: G. Androutsopoulos, Youth language. In A.F. Christidis (eds.), *Encyclopedic guide for the language* (pp. 108–113). Thessaloniki: Centre of Greek Language.
- Baldwin & Kim 2010: T. Baldwin, S.N. Kim, Multiword expressions. *Handbook of natural language processing*, 2, 267–292.
- Bella 2015: S. Bella, *Pragmatics. From language communication to language teaching*. Athens: Gutenberg.
- Cacciari & Levorato 1998: C. Cacciari, M.C. Levorato, The effect of semantic analyzability of idioms in metalinguistic tasks. *Metaphor and Symbol*, 13(3), 159–177.
- Calzolari et al. 2002: N. Calzolari, C.J. Fillmore, R. Grishman, N. Ide, A. Lenci, C. MacLeod, A. Zampolli, Towards Best Practice for Multiword Expressions in Computational Lexicons. In *LREC*, vol. 2, (pp. 1934–1940).
- Constant et al. 2017: M. Constant, G. Eryiğit, J. Monti, L. Van Der Plas, C. Ramisch, M. Rosner, A. Todirascu, Multiword expression processing: A survey. *Computational Linguistics*, 43(4), 837–892.
- Fotopoulou 1993: A. Fotopoulou, *Une classification des phrases à compléments figés en grec moderne - étude morphosyntaxique des phrases figées* [PhD Thesis]. Paris: University Paris VIII-St Denis.
- Fotopoulou & Giouli 2018: A. Fotopoulou & V. Giouli, MWEs and the Emotion Lexicon: typological and cross-lingual considerations. In M. Seiler & S. Markantonatou (eds). *Multiword Expressions: Insights from a Multi-lingual Perspective*. Berlin: Language Science Press series.

- Fotopoulou, Laporte & Nakamura 2021: A. Fotopoulou, E. Laporte, T. Nakamura, Where Do Aspectual Variants of Light Verb Constructions Belong?. In *Proceedings of the 17th Workshop on Multiword Expressions (MWE 2021)* (pp. 2–12).
- Gibbs et al. 1989: R.W. Gibbs, N.P. Nayak, C. Cutting, How to kick the bucket and not decompose: Analyzability and idiom processing. *Journal of memory and language*, 28(5), 576–593.
- Gross 1984: M. Gross, Lexicon-Grammar and the syntactic analysis of French. *International Conference on Computational Linguistics (COLING)* (pp. 275–282). Stanford: Association for Computational Linguistics.
- Gross 1986: M. Gross, Lexicon-grammar. In *Lexicon-Grammar. The Representation of Compound Words* (pp. 1–6).
- Gross 1996: G. Gross, *Les expressions figées en français: Noms composés et autres locutions*. Paris: Ophrys.
- Harris 1954: Z.S. Harris, Distributional structure. *Word*, 10(2–3), 146–162.
- Harris 1970: Z.S. Harris, Transformational theory. In *Papers in Structural and Transformational Linguistics*, vol 1, (pp. 533–577). Dordrecht: Springer.
- Katos 2016: G. Katos, *Dictionary of our Vernacular and Marginal language*, Centre of Greek Language [Online], [<http://georgakas.lit.auth.gr/dictionaries/index.php/leksika/katou-g>].
- Kyriazi 2020: P. Kyriazi, *Youth Language: expression of identity via virtual communication* [Unpublished undergraduate thesis]. Athens: National and Kapodistrian University of Athens.
- Kyriacopoulou & Sfetsiou 2002: T. Kyriacopoulou, V. Sfetsiou, Les constructions nominales à verbe support en grec moderne. *Linguistic Insights, Phrases and Phraseology – Data and Descriptions*, 5, 163–181.
- Lamiroy 2003: B. Lamiroy, Grammaticalisation et comparaison de langues. *Verbum*, 25(3), 411–431.
- Mavrogiorgos 2010: M. Mavrogiorgos, *Clitics in Greek: a minimalist account of proclisis and enclisis*. Amsterdam: John Benjamins Publishing Company.
- Markantonatou & Samaridi 2018: S. Markantonatou, N. Samaridi, Verb MWEs and the grammatical function ‘object’. In M. Seiler & S. Markantonatou (eds). *Multiword Expressions: Insights from a Multi-lingual Perspective*. Berlin: Language Science Press series.
- Mini 2009: M. Mini, *Linguistic and psycholinguistic study of fixed verbal expressions with fixed subject in Modern Greek: morphosyntactic analysis, semantic gradation and processing by children of school age*. [Phd Thesis], Patra: University of Patras.
- Moustaki 1995: A. Moustaki, *Les expressions figées είμαι/être Prép C W en grec moderne*. [PhD Thesis]. Paris: Université Paris VIII.

- Nunberg 1994: G. Nunberg, I. A. Sag, T. Wasow, Idioms. *Language*, 70(3), 491–538.
- Philippaki-Warbuton 1994: I. Philippaki-Warbuton, Verb movement and clitics in Modern Greek. *Amsterdam Studies in The Theory and History of Linguistic Science Series 4*, 53–53.
- Philippaki-Warbuton *et al.* 2002: I. Philippaki-Warbuton, S. Varlokosta, M. Georgiafentis, G. Kotzoglou, On the status of ‘clitics’ and their ‘doubles’ in Greek. *Reading Working Papers in Linguistics*, 6, 57–84.
- Ralli & Stavrou 1998: A. Ralli, M. Stavrou, Morphology-syntax interface: AN compounds vs. AN constructs in Modern Greek. In *Yearbook of morphology 1997* (pp. 243–264). Dordrecht: Springer.
- Ralli 2005: A. Ralli, *Morphology*. Athens: Patakis Publications [In Greek].
- Roussou 2015: A. Roussou, *Syntax*. Athens: Link of Hellenic Academic Libraries.
- Sag *et al.* 2002: I. A. Sag, T. Baldwin, F. Bond, A. Copestake, D. Flickinger, Multiword expressions: A pain in the neck for NLP. In *Computational Linguistics and Intelligent Text Processing: Third International Conference, CICLing 2002 Mexico City, Mexico, February 17–23, 2002 Proceedings 3* (pp. 1–15). Berlin: Springer.
- Spencer & Luis 2012: A. Spencer, A.R. Luís, *Clitics: an introduction*. Cambridge: Cambridge University Press.
- Tzartanos 1930: A. Tzartanos, *Grammar of Modern Greek*. Athens: Kakoulidhis.
- Xydopoulos 2008: G.I. Xydopoulos, *Lexicology: Introduction to word and lexicon analysis*. Athens: Patakis Publications.

Παναγιώτα Κυριαζή
 Ινστιτούτο Επεξεργασίας του Λόγου (ΙΕΛ), Ερευνητικό Κέντρο “Αθηνά”, Ελλάδα
 Αγγελική Φωτοπούλου
 Ινστιτούτο Επεξεργασίας του Λόγου (ΙΕΛ), Ερευνητικό Κέντρο “Αθηνά”, Ελλάδα

**ΠΟΛΥΛΕΚΤΙΚΕΣ ΕΚΦΡΑΣΕΙΣ ΤΗΣ ΕΛΛΗΝΙΚΗΣ ΓΛΩΣΣΑΣ: ΜΙΑ ΜΕΛΕΤΗ ΓΙΑ ΤΑ ΜΗ
 ΑΝΑΦΟΡΙΚΑ ΚΛΙΤΙΚΑ ΣΤΙΣ ΠΟΛΥΛΕΚΤΙΚΕΣ ΕΚΦΡΑΣΕΙΣ**

Περίληψη

Σκοπός της εν λόγω εργασίας αποτελεί η παρουσίαση δύο διαφορετικών ταξινομήσεων των κλιτικών εντός ρηματικών πολυλεκτικών εκφράσεων (ΠΛΕ) της Νέας Ελληνικής, τα οποία φαίνονται εκ πρώτης όψεως να μην έχουν συγκεκριμένη αναφορά. Οι ταξινομήσεις πραγματοποιήθηκαν επί τη βάσει διαφορετικών κριτηρίων: η πρώτη ταξινόμηση βασίζεται στη συντακτική δομή των ρηματικών εκφράσεων σύμφωνα με τη θεωρία του Λεξικού-Γραμματικής του Maurice Gross και η δεύτερη δομείται ανάλογα με το βαθμό αναλυσιμότητας των κλιτικών. Μέσα από διεξοδική ανάλυση και εφαρμογή διαγνωστικών ελέγχων (συντακτικά και λεξικά κριτήρια) 132 σωμάτων κειμένων με ρηματικές ΠΛΕ, δομούνται δύο βασικές κατηγορίες: τα μη-αναλύσιμα και αναλύσιμα κλιτικά με συνολικά 5 υποκατηγορίες (2 στην πρώτη και 3 στη δεύτερη). Παρατηρούμε τα κλιτικά να συμπεριφέρονται άλλοτε ελεύθερα εντός ημι-παγιωμένων εκφράσεων και άλλοτε παγιωμένα εντός στερεότυπων δομών σε συνάρτηση με τη δυναμική σχέση που αναπτύσσει η προρηματική αντωνυμία με την ρηματική φράση (ΡΦ). Κατ’ αυτόν τον τρόπο, η αναλυσιμότητα των κλιτικών εκτείνεται από την εντελώς παγιωμένη δίχως αναφορά και ανάλυση στην πιο ελεύθερη με τα κλιτικά να αναλύονται και να αναφέρονται σε ονοματικές φράσεις εντός του λόγου. Ο βαθμός της αναλυσιμότητας των προρηματικών αντωνυμιών εντός των ρηματικών εκφράσεων περιγράφεται ως το *συνεχές της αναλυσιμότητας των κλιτικών* κατ’ αναλογία προς τον ορισμό του G. Gross (1996), το *συνεχές της παγίωσης*, ο οποίος περιγράφει το βαθμό της αναλυσιμότητας των ΠΛΕ από τις απόλυτα στερεότυπες στις ημι-παγιωμένες και, τέλος, στις επιβοηθητικές ρηματικές δομές σύμφωνα με τις κατανομικές και σημασιολογικές ιδιότητες των συστατικών στοιχείων της εκάστοτε έκφρασης.

Λέξεις-κλειδιά: πολυλεκτικές εκφράσεις, συνεχές της αναλυσιμότητας, κλιτικά, βαθμός παγίωσης

Appendices

Appendix 1

Symbols	Syntactic/transformational properties
Det	Determiner of any kind
Adj	Adjective
Adv	Adverb
Prep	Preposition
N, C	Non-fixed or fixed noun phrase or noun. The numerical indexes next to N or C refer to their syntactic role in the clause
Nhum	Noun phrase necessarily animate
N-hum	Noun phrase necessarily inanimate
Npc	Noun phrase indicating parts of the body
N(gen)	Noun in genitive case
QuP	Subordinate clause
Ppv	Personal pronoun (inflectional) inserted before the verb
Poss	Possessive pronoun. There may be numerical indexes like 1,2,3 indicating the required co-reference with a component of the sentence: Poss0, Poss1, Poss2...
V	Verb
Vsup	Support verb
E	Empty set
Negation	Negation
Passif	Transformation of passivisation
Perfect Participle	Perfect Participle
Permutation	Permutation
Symetrie	Symmetry
Active	Active voice
Effacement	Effacement
+	Symbol in tables indicating that the property written in the headline of the column is valid for the verb
-	Symbol in tables indicating that the property written in the headline of the column is invalid for the verb
=:	Symbol used to specify the lexical or syntactic content of a structure (rewriting)

Appendix 2
N0 PPV V C1

N0 =: Nhum	N0 =: N-hum	Negation	Clitic	V	Det	C1	C1 =: Npc	C1 =: Adv	Phrase Vsup	Symetrie	Examples	Comments
+	-	δεν	το	βάζω	-	κάτω	-	+	-	+	Μην το βάζεις κάτω...Θα σου γελιάσει η τύχη	
+	-		τα	βάφω	-	μαύρα	-	-	-	-	Δεν θα τα βάψουμε και μαύρα επειδή δεν ήρθε, σινά! Την έβγαλε λάδι που δεν χρειάστηκε να πάει και σε δεύτερο γιάτρο	Συνηθως συναντάται σε χρόνο παρελθοντικό
+	-		την	βγάζω	-	λάδι	-	-	-	+	Τελικά πάλι καθαρή την έβγαλες...	Συνηθως συναντάται σε χρόνο παρελθοντικό
+	-		την	βλέπω	-	μαύρα κι άραχνα	-	+	+	+	Αν τα βλέπεις μαύρα κι άραχνα δεν θα κατάφερες και τίποτα	
+	-		τα	βλέπω	-	ρόδινα	-	+	+	+	Ο Δημήτρης είναι μερικές φορές πολύ αφελής, τα βλέπει όλα ρόδινα	
+	-		τα	έχω	-	τετρακάσια	-	-	-	-	Θέλω να γεράσω, να γίνω μπαμπούρια, να τα έχω τετρακάσια και να περπατώ	Ενν. τα μυαλά
+	-		την	έχω δει	-	κάποιος	-	-	-	-	Την έχει δει κάποιος τώρα τελευταία	Συναντάται σε παρακείμενο και αόριστο
+	-		την	κάνω	-	λαχείο	-	-	-	-	Καλά τόσο ωραίο ξεχονοχείο σε τόσο χαμηλή τιμή; Την κάναμε λαχείο!	Συνηθως σε χρόνο αόριστο
+	-		το	βάζω	-	κάτω	-	+	-	+	Δεν θα τα βάψουμε και μαύρα επειδή δεν ήρθε, σινά!	