

John Benjamins Publishing Company



This is a contribution from *Diachronica* 35:3
© 2018. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's/s' institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu.

Please see our rights policy on <https://benjamins.com/content/customers/rights>

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Non-configurationality in diachrony

Correlations in local and global networks of Ancient Greek and Latin

Edoardo Maria Ponti and Silvia Luraghi

University of Cambridge / University of Pavia

Non-configurationality is a linguistic property associated with free word order, discontinuous constituents, including NPs, and null anaphora of referential arguments. Quantitative metrics, based both on local networks (syntactic trees and word order within sentences) and on global networks (incorporating the relations within a whole treebank into a shared graph), can reveal correlations among these features. Using treebanks we focus on diachronic varieties of Ancient Greek and Latin, in which non-configurationality tapered off over time, leading to the largely configurational nature of the Romance languages and of Modern Greek. A property of global networks (density of their spectra around zero eigenvalues) measuring the regularity in word order is shown to be strengthened from classical to late varieties. Discontinuous NPs are traced by counting the words creating non-projectivity in dependency trees: these drop dramatically in late varieties. Finally, developments in the use of null referential direct objects are gauged by assessing the percentage of third-person personal pronouns among verb objects. All three features turn out to change over time due to the decay of non-configurationality. Evaluation of the strength of their pairwise correlation shows that null direct objects and discontinuous NPs are deeply intertwined.

Keywords: non-configurationality, treebanks, network analysis, non-projectivity, discontinuous constituents, referential null objects

1. Introduction

In this paper, we show how treebank-based queries and network analysis allow us to measure the development of a number of features of Classical Greek and Latin syntax that are normally considered correlates of non-configurationality, that is, free constituent order, discontinuous NPs and use of null anaphora for definite referential

direct objects. We chose these features as diagnostics for non-configurationality and their decay as a hint to the rise of configurationality based on Baker (2001: 1434), who writes: “[i]n the narrow sense, a nonconfigurational language is one that has ... free word order, possible omission of all grammatical functions, and the possibility of having discontinuous NP constituents.”

While Classical Greek and Latin displayed these features, both Modern Greek and the Romance languages feature configurational syntax to a large extent. We aim to capture the ongoing rise of configurationality based on two diachronic treebanks of Ancient Greek and of Latin available from the PROIEL project (see §3.1). The analysis is based on quantitative parameters associated with features of non-configurationality, and these allegedly co-vary in time. They are measured both at the local and at the global level (syntactic trees and co-occurrences in single sentences and networks; see §3.2). Our paper is organized as follows. In §2 we discuss the notion of non-configurationality and how it applies to Classical Greek and Latin. In §3 we describe the experiment setup and the data. Section 4 is devoted to the formal definition of individual metrics related to non-configurationality and the assessment of their values. In §5 we present the analysis of these results, and finally we draw some general conclusions in §6.

2. Non-configurationality

The term ‘non-configurationality’ was introduced in Hale’s (1983) study of Warlpiri, in order to account for a number of typical features of this language that make it remarkably different from languages like English. According to Hale (1983), non-configurational languages have a ‘flat’ structure, or a hierarchical structure at the level of Lexical Structure only, which does not project on Phrase Structure. This observation leads to the conclusion that the VP is not relevant in non-configurational languages, in which, typically, “subjects and objects cannot be identified by word order and simple constituency tests in any straightforward way” (Baker 2001: 1433).

Research on non-configurationality first developed within the Government and Binding framework, but in recent years, as features of non-configurationality have been reported from numerous languages of different genetic and areal affiliations, it has increasingly attracted the interest of typologists (for a survey, see Reinöhl 2016: 23–27, 45–48). As we remarked in §1, typical correlates of non-configurational languages have been shown to be free (i.e., pragmatically determined) word order, discontinuous NPs and extensive null realization of definite referential arguments even when they are not co-referenced on the verb (Austin & Bresnan 1996; Baker 2001).

Current research shows that configurationality should be regarded as a gradient property, as languages may be configurational or non-configurational to different extents. For example, Hungarian has been argued to be non-configurational in clause structure, as it allows free constituent order, but configurational in noun phrase structure, as it does not allow discontinuous NPs (Kiss 1987). Indeed, free constituent order is not necessarily associated with a high degree of non-configurationality: the fact itself that constituency is relevant at the phrasal level implies, for example, that discontinuous NPs are normally not allowed in languages such as Hungarian. We return to this issue in §2.1.

Configurationality can arise as a result of language change and become more extensive over the course of time. Ancient Indo-European languages show typical features of non-configurationality, including among other things free word order, discontinuous NPs and definite referential null objects which are not co-indexed on the verb (Devine & Stephens 2000; Schäufele 1990; Rögnvaldsson 1995; Luraghi 1997, 2003), a weak noun-adjective distinction (see §2.2) and the trend toward increasing configurationality has been described for many of them (Luraghi 2010). As Reinöhl (2016: 45) remarks “Latin and Greek only possess incipient phrasal structures, ... (Latin having prepositional phrases and Ancient Greek developing nominal expressions involving articles), ... Vedic shows a lack of such structures.” According to Hewson & Bubenik (2006), configurationality in Indo-European languages first manifested itself with the increasing grammaticalization of adpositional phrases and the creation of adpositions out of earlier adverbs. Reinöhl (2016), though distancing her views in relevant respects from those of Hewson and Bubenik, also argues that the rise of adpositions brought about configurationality. Ledgeway (2012) shows that the prepositional phrase and, while less developed, the complementizer phrase already existed in Latin. Indeed, adpositional phrases are fully grammaticalized not only in Latin, but also in Classical Greek; for this reason we do not take them into account.

In the following sections, we show how certain correlates of non-configurationality are instantiated in Classical Greek and Latin.

2.1 Word order

Classical Greek and Latin are so-called free word order languages. The position of the verb in the sentence is sensitive to pragmatic factors: it may show author-specific preferences, but it is not restricted from occurring in sentence initial, internal or final position. In particular, concerning the position of the verb in Herodotus’ *Histories*, which constitute part of the corpus for this paper, a partial analysis carried out by Dover (1960) yields the following counts for word order patterns: VS (113) vs SV (174); VO (203) vs OV (161), with a preference for post-verbal direct

objects. Other authors analyzed by Dover include Lysias, who shows a preference for pre-verbal direct objects and final verbs, and Plato, who has approximately the same percentage of OV and VO occurrences.

Latin is often referred to as an SOV language, and final verbs do in fact predominate in all authors. However, initial and internal position are also possible options in all literary genres and at all diachronic stages. Caesar, dubbed a ‘final position fanatic’ (*Fanatiker der Endstellung*) by Linde (1923: 154), has the verb in final position in 84% of main clauses and 93% of subordinate clauses. For Cicero, Linde (1923: 155) found around 50% of final and 50% of non-final verbs in main clauses, with variation among different types of work and a considerably higher proportion of final verbs in subordinate clauses. Similarly, Danckaert (2015: 241) in a survey of various studies of word order in Cicero’s works, signals a range of variation from 63.1% to 95.9% in OV sentences (the figures cover both main and subordinate clauses).

Though considered SVO languages, both Modern Greek and the Romance languages allow free constituent order to a varying extent, partly due to extensive use of direct object clitics, which are usually preverbal and can co-index displaced constituents. Changes in clause structure have been observed over the history of these languages (Deligianni 2011; Revithiadou & Spyropoulos 2007, 2008; Salvi 2004; Luraghi 1998, 2010; Ledgeway 2011, 2012). Constraints on the order of constituents necessarily follow the rise of constituency. For this reason, constituent order can be diagnostic for ongoing change from Classical to late varieties of Greek and Latin only if connected with the decay of other correlates of non-configurationality, such as null objects and discontinuous NPs.

2.2 Discontinuous NPs

According to Baker (2001: 1437), discontinuous NPs “are possible only in languages with no more than a weak N[oun]/A[djective] contrast,” because syntactically adjectives are predicates of nouns, rather than being dependent. This is similar to the traditional view on Proto Indo-European adjectives, as expressed for example by Meillet & Vendryes (1924: 530): “Adjectives are by no means connected with nouns. They are usually inflected in the same case, same number, and, as distinctive for adjectives, same gender ... because they refer to the same entity.”¹ Discontinuous constituents in classical varieties of Latin and Greek occur to varying extents depending on literary genres but are well attested in literary prose.

1. L’adjectif n’est nullement lié au substantif. Il est généralement au même cas, au même nombre, et, ce qui est le trait caractéristique de l’adjectif, au même genre ..., mais parce qu’il s’applique au même objet.

Examples of discontinuous NPs from Latin are (1) and (2); see further *duas legiones ... novas* in (7).

- (1) *aliquo te cum hoc rei publicae vinculo esse coniunctum*
 INEF.ABL 2SG.ACC with DEM.ABL state:GEN link:ABL bind:INEF.PF.P
 “(that) you were bound to him by some responsibility for the state.”
 (Cic. *Mur.* 64)

- (2) *neque quisquam agri modum certum aut fines*
 nor INDEF.NOM land:GEN measure:ACC certain:ACC or border:ACC.PL
habet proprios
 have:PRS.3SG own:ACC.PL
 “Nor has anyone a fixed quantity of land or his own individual limits.”
 (Caes. *Gal.* 6.22.2)

Classical Greek NPs differ from Latin mainly due to the existence of fully grammaticalized definite articles. This should point toward a higher degree of configurationality. Note however that even NPs with definite articles allow for various types of discontinuity (see Devine & Stephens 2000 for an exhaustive description). In particular, occurrences in which the definite article is separated from the noun it determines, when a constituent is sentence initial, as in (3), should receive separate treatment.

- (3) *ho de khrosós hoútos kai ho árguros*
 ART.NOM PTC gold:NOM DEM.NOM and ART.NOM silver:NOM
kaléetai Gygádas
 call:PRS.M/P.3SG Gygian:NOM
 “This gold and the silver are called Gygian.” (Hdt. 1.14.3)

Notably, items that can stand between the definite article and the noun are so-called postpositives, that is, second position, or P2, particles that may bear a graphic accent but prosodically behave as clitics. We return to this issue in §4.1.

Often, discontinuity is caused by the occurrence of a clitic, as in (4). In this example, the direct object clitic *min* is not only separated from the verb *apopémpseie* “had sent”, it also splits up the NP *hoíon ándra* “such man”. Notably, *min* is often described as a P2 clitic. However, in Classical Greek especially, pronominal P2 clitics could be placed elsewhere in the sentence, as described in Goldstein (2016), even though they did not show any special preference for a specific type of constituent; see further Luraghi (2013).

- (4) *thōmázein te autoû par' hoíōn min ándra*
 wonder:INF.PRS PTC there by such:ACC 3SG.ACC man:ACC
apopémpseie
 send:OPT.AOR.3SG
 “(The herald) wondered what sort of man he had been sent to.” (Hdt 5.92f3)

The occurrence of discontinuous NPs is a major difference between Classical Greek and Latin on the one hand and Modern Greek and Romance on the other. As the data in §4.1 show, the number of discontinuous NPs drops by 86.2% from Classical to Late Greek and by 89.48% from Classical to Late Latin. One of the few discontinuous NPs in the Late Greek corpus is *héteron doûlon* “another servant” in (5a). Interestingly, the Latin translation in (5b) does not mirror the same discontinuity. Modern Greek in (5c) and Italian in (5d) also contain continuous NPs.

- (5) a. *kai prosétheto héteron pémpsai doulon*
 and add:AOR.MID.3SG other:ACC send:INF.AOR servant:ACC
 “He sent yet another servant.” (Luke 20.11)
 b. *Et addidit alterum servum mittere*
 c. *Apéstile ke páli énan állon dulo*
 d. *Mandò un altro servo*

2.3 Definite referential null objects

Null arguments are common in Latin and in Classical Greek. Both languages make extensive use of null subjects; however, as subjects are extensively co-indexed on finite verbs through a complex morphological system of agreement, the occurrence of null subjects is not indicative of non-configurationality.

Much more significant is the occurrence of null referential direct objects, as direct objects are not co-indexed on the verb. Null referential direct objects occur in different syntactic and discourse conditions, as in Examples (6) and (7) (see Luraghi 1997, 2003; Keidana & Luraghi 2012).

- (6) *Epexêlthon hoí te epíkouroi kai autôn*
 march:AOR.3PL ART.NOM.PL PTC mercenary:NOM.PL and DEM.GEN.PL
Samiôn sukhnoí, dexàmenoí dè toûs
 Samian:GEN.PL many:NOM.PL engage:PTCP.AOR.NOM.PL PTC ART.ACC.PL
Lakedaimoníous ep' olígon khrónon épheugon opísō
 Spartan:ACC.PL on little:ACC time:acc flee:IMPF.3.PL back
hoí dè epispómenoí Ø_i ékteinon Ø_i
 ART.NOM.PL PTC pursue:PTCP.AOR.MID.NOM.PL kill:IMPF.3PL

“The mercenaries and many of the Samians themselves sallied out near the upper tower on the ridge of the hill and withstood the Lacedaemonian advance for a little while; then they fled back. The Lacedaemonians pursuing **them** destroyed **them**.” (Hdt. 3.54.2)

- (7) *Caesar duas legiones, in citeriore Gallia novas*
 Caesar:NOM two:ACC legionACC.PL in hither:ABL Gaul:ABL new:ACC.PL
conscriptis et inita aestate, in interiorum
 enroll:PF.3SG and begin:PTCP.PF.ABL summer:ABL in inner:ACC
Galliam qui Ø_i deduceret, Quintum Pedium
 Gaul:ACC REL.NOM lead:SBJ.IMPF.3SG Quintus:ACC Pedius:ACC
legatum misit
 lieutenant:ACC send:PF.3SG

“Caesar enrolled two new legions in Hither Gaul and at the beginning of the summer he sent Quintus Pedius, lieutenant-general, to lead **them** into Inner Gaul.” (Caes. G. 2.2)

In particular, in cases in which the same direct object is shared by two coordinated clauses (object sharing), deletion in the second clause seems to be mandatory. In Latin, no exceptions have been found in Classical authors. The only occurrences in which a direct object can be repeated in coordinated clauses appear in contexts in which more than one possible antecedent is available, as in (8), or in which the direct object in the second conjunct is accented for emphasis, as in (9). In the latter case, the emphatic pronoun typically hosts the clitic conjunction =*que*. Occurrences similar to these from other Latin authors are thoroughly discussed in Luraghi (1997).

- (8) *litteras scripsi hora decima Cerialibus*
 letter:ACC.PL write:PF.1SG hour:ABL tenth:ABL Cerealia:ABL.PL
statim ut tuas legeram sed eas
 immediately as POSS.2PL.ACC.PL read.PPF.1SG but 3SG.ACC.PL
eram daturus ut putaram postridie
 be:IMPF.1SG give:PTCP.FUT.NOM as think:PPF.1SG next.day
 “I wrote (you) a **letter** at four o’clock in the afternoon of the Cerealia as soon as I received yours [possible conflicting antecedent], and I was thinking of giving it the next day (to the first available person).” (Cic. Att. 2.12.4)

- (9) *postero autem die Caesar ... Vettium in rostra*
 next:ABL however day:ABL Caesar:NOM Vettius:ACC in roster:ACC.PL
produxit eum= que in eo loco constituit quo
 bring:PF.3SG 3SG.ACC and in DEM.ABL place:ABL place:PF.3SG where
Bibulo consuli adspirare non liceret
 Bibulus:DAT consul:DAT hope:INF NEG be.allowed:SBJ.IMPF.3SG

“However, the next day Caesar took Vettius on the rostra and placed him in a position in which Bibulus, though being consul, was not allowed to stand.”

(Cic. Att. 2.24.3)

In the case of direct object sharing, Classical Greek too mostly features pronouns in the second coordinated clause when the context contains more than one possible antecedent (see Luraghi 2003; Keydana & Luraghi 2012). However, in Herodotus one also finds at least one exception, discussed in Keydana & Luraghi (2012: 119), featuring overt realization of a pronominal direct object in the second conjunct and the coordinating conjunction *kaí*. In (10), the direct object is overtly realized in the second conjunct, featuring the adversative particle *allá*. In fact, in this passage the adversative character of the second coordinate clause may have favored repetition of the direct object (note further that the statue is the topic of a long stretch of discourse in the preceding context).

- (10) *tòn dè andriánta tou̐ton Délíoi ouk apégagon*
 ART.ACC PTC statue:ACC DEM.ACC Delian:NOM.PL NEG remove:AOR.3PL
allá min di' etéōn eíkosi Thēbaíoi ... ekomísanto
 but 3SG.ACC for year:GEN.PL twenty Theban:NOM.PL bring:AOR.MID.3PL
epi Délion
 toward Delion:ACC
 “But the Delians never carried **that statue** away; twenty years later the Thebans brought it to Delium.”
 (Hdt. 6.118.3)

Referential null objects disappeared with the rise of configurationality in Romance (Luraghi 1998; Ledgeway 2012) and Modern Greek (Revithiadou & Spyropoulos 2007, 2008), which developed a system of pronominal clitics.

3. Methodology

We use a treebank corpus to explore non-configurationality in Latin and Ancient Greek. As mentioned in §1, we observe the relevant features of non-configurationality on both the local and the global levels of linguistic networks, following the distinction proposed by Čech et al. (2011). The local level consists of the syntactic dependency trees and word order of the individual sentences (see §3.1). The global level consists of a single network constructed from a treebank with a technique pioneered by Ferrer i Cancho & Solé (2001): each distinct lemma corresponds to a node, whereas each distinct relation between a pair of lemmas corresponds to an edge directed from one lemma to the other. This relation can be either co-occurrence, meaning that a word follows another in the linear order of a sentence, or dependency, meaning that a word is the parent of another in a syntactic tree. Global

networks mirror holistic properties of a language, possibly different from the sum of the properties of the local networks (Solé et al. 2010; Baronchelli et al. 2013). Local networks in turn are better suited for identifying fine-grained phenomena. As a consequence, both levels are necessary to capture non-configurationality, as it affects both a language variety as a whole and some of its specific constructions. In the rest of this section, we describe the data from Latin and Ancient Greek that we used for this paper, and we outline the method used to generate co-occurrence and dependency networks from the data.

3.1 The corpus

The data come from the collection of dependency treebanks developed within the PROIEL project (*Pragmatic resources in old Indo-European languages*; see Haug & Jøhndal 2008). A dependency treebank is a corpus of texts annotated with dependencies at the syntactic layer. Sentences are represented as trees where each word corresponds to a node: top-down relations indicated by edges convey grammatical relations between a head and a dependent. In addition, nodes are arranged on the left-to-right dimension to convey the linear order, i.e., the precedence relations.

PROIEL contains treebanks for several ancient Indo-European languages and for different varieties of the same language. We selected four of these treebanks, Ancient Greek and Latin, both Classical and Late. The amount of tokens was equalized to the count of the smallest treebank (67,247 tokens) approximated to the closest sentence boundary. The four treebanks consist of the texts listed in Table 1 (in parentheses we indicate the actual span of the text used for this work).

Table 1. Texts composing the treebanks

	Classical Greek	Late Greek	Classical Latin	Late Latin
Author	Herodotus	Septuagint	Caesar and Cicero	Jerome
Title	<i>Histories</i> (I.1–VII.83)	<i>New Testament</i> (Matthew I.1 – Acts of the Apostles V.10)	<i>The Gallic War</i> (I.1– VII.77) and <i>Letters to</i> <i>Atticus</i> (I.1–VI.9)	<i>Vulgate</i> (Genesis I.1 – Acts of the Apostles XIV.11)
Date	440–429 BCE	49–150 ca. CE	58–50 BCE and 68–43 BCE	382–413 CE

We preferred the PROIEL treebanks over the Perseus collection (Bamman et al. 2009) for a series of reasons. In the first place, the two sets of treebanks cannot be merged, as they rely on incompatible annotation schemes. Also, Ancient Greek texts in the Perseus treebank are mostly poetry, and further, the Latin texts are limited in size. Hence we consider the texts in the PROIEL treebank the best approximation

available for the relevant language varieties, even though we are aware of the fact that they are not entirely representative, as they are mostly limited to single authors, and more variables than just diachrony separate them, notably social status of the authors and literary genre of the texts.

3.2 Network induction

A network is a graph consisting of a set of nodes V and a set of edges E . A network can be induced from a treebank by setting up an equivalence between (i) nodes and properties of words (e.g., their form, lemma, part-of-speech tag, etc.) and (ii) edges and word relations (e.g., precedence in linear order, dependency in syntax, etc.). Ferrer i Cancho et al. (2004) developed a method to create networks based on lemmas as nodes and dependencies as edges. Networks are useful because they mirror global properties of a language that can be hidden in local structures (the dependency trees). For example, linguistic networks of children's speech show a sudden change from tree-like structures to scale-free, small-world structures (i.e., hierarchical and highly connected) around the age of two years (Solé et al. 2010). This happens simultaneously with the appearance of functional words and inflectional morphology. Furthermore, linguistic networks cast light on the nature of linguistic universals and variation. Čech et al. (2011), for example, demonstrated that verbs behave as hubs (i.e., nodes with many connections) in linguistic networks cross-linguistically, providing evidence in support of Tesnière's predicate-centric theory (Tesnière 1959).

Although recent work on linguistic networks has focused on dependency networks, these completely obscure another range of properties of languages: those affecting the linear order of words. Free word order (including the fluctuation in the verb position) is such a property. Hence, in this work we explore linguistic networks based on co-occurrence and dependencies in order to investigate to what extent they diverge. Thus far, networks based on linear order have exploited collocations, i.e., co-occurrences more frequent than chance (Ferrer i Cancho & Solé 2001; Kapustin & Jansen 2007), as an approximation of dependency-based networks. Our method does not filter out any co-occurrence: each lemma corresponds to a node, whereas an edge is created between every two adjacent lemma instances for co-occurrence networks and between every head-dependent pair for dependency networks. The edges of these graphs are oriented, in order to distinguish right and left (for linear order) or top and bottom (for syntactic dependencies) contexts of the nodes, and loops (edges departing from and arriving to the same node) are forbidden.

Figure 1 displays the dependency tree (left) and the equivalent lemma-based co-occurrence network (center) of Example (11). The latter is constructed by creating a node for each distinct lemma of the words in the tree (e.g., *et* for *et* and *voco*

for *vocabis*). Then a directed edge is created between two nodes if one immediately follows the other in the tree (e.g., *et* and *voco*). Repeating this procedure over the whole treebank of Late Latin results in a global network (right). For the sake of visualization, this is shown in such a way that the more an edge tends to the center, the more frequent it is (note that this choice of display has no effect on the properties of the network). These graphs can be equivalently specified by a binary adjacency matrix *A*. Each row and each column of the matrix corresponds to a separate node. For a node pair *i* and *j*, the cell $A_{i,j}$ is filled with 1 if there exists an edge between them, otherwise it has 0.

(11) *Et vocabis nomen eius Iesum*
and call:FUT.2SG name:ACC 3SG.GEN Jesus:ACC
“And you will call him Jesus.” (Matthew 1.21)

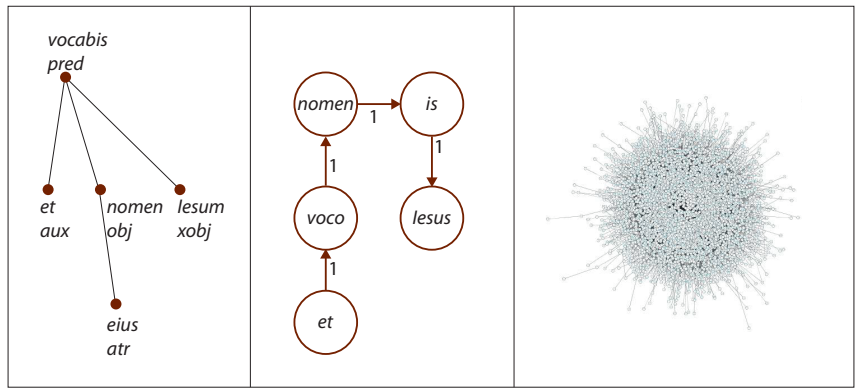


Figure 1. An example of syntactic tree (left), its equivalent network representation (center) and the global network resulting from a similar transformation of the whole treebank

We created both co-occurrence and dependency networks of lemmas for each treebank. Basic information about them is summarized in Table 2, which reports the total number of nodes and edges. Note that the late varieties have lower figures for both, because of inherent properties of the texts. In particular, for the same number of tokens, they have a smaller set of lemmas, which on average have a higher frequency compared to the set of classical varieties.

Table 2. Number of nodes and edges of the induced networks

Name	Nodes (lemmas)	Edges
Classical Greek	5398	34076
Late Greek	3025	20788
Classical Latin	4824	39311
Late Latin	3071	25021

4. Metrics and results

As we argued in §2, non-configurationality has been related to free word order, discontinuous NPs and null anaphora of definite referential direct object. In this section, we propose some metrics to assess quantitatively both the pervasiveness of these phenomena in a language variety and the difference between two diachronic stages. These metrics hinge on properties of global networks or queries constrained by word order and syntactic dependencies in local networks.

4.1 Free word order

The feature of unconstrained word order is problematic for many reasons. In the first place, it surfaces in languages that are not, strictly speaking, non-configurational (Luraghi 2010). In addition, neither a clear formal definition nor any sound method to measure this feature is available. Futtrel et al. (2015) propose a measure of freedom of word order (i.e., argument order with respect to the verb) based on conditional entropy, that is, the uncertainty in determining a word sequence given an unordered dependency tree. They demonstrate that classical varieties of Ancient Greek and Latin are the languages with the highest entropy in terms of word order and head-dependent directionality among the treebanks in the Universal Dependencies collection (Nivre et al. 2016), including Modern Greek and Romance languages. However, this measure turns out to be unreliable if ranging over all the syntactic relations because of the difficulties in estimating entropy statistically and avoiding data sparsity due to the long-tail distribution of linguistic phenomena.

With respect to the position of verbs, we counted whether objects depending on a verb follow it (VO) or precede it (OV) in Table 3, obtaining results in line with the expectations of §2.1. For Ancient Greek, we observe a shift from indifference regarding OV or VO order to a clear preference for VO. In Latin, the preference changes by swapping the order from OV to VO.

Table 3. Counts of objects following (VO) or preceding (OV) verbs in the four treebanks

	VO	OV
Classical Greek	3762	3409
Late Greek	4501	2158
Classical Latin	1473	5471
Late Latin	4884	2791

Moreover, following Ponti (2016), we propose as an alternative identifying across-the-board word order freedom with the ‘irregularity’ of a global network.

The gist is that, if a lemma is allowed to appear in more contexts, then its neighborhood in the network is more idiosyncratic and does not match the neighborhood of similar words. For instance, if a verb like *fero* usually appears before a set of object nouns, they will be linked together in the network. If another verb like *accipio* shows similar behavior, then its neighbors will overlap (at least in part) with *fero*. However, if both verbs can occur in any position in the sentence, no syntactic regularity forbids the neighbors to be different.

This measure of irregularity is more reliable than other topological properties of networks such as Clustering Coefficient or Average Minimum Path Length, because those properties may be skewed by the size of the network. Indeed, on account of idiosyncratic properties of the texts, both networks of the late varieties appear to have a smaller number of nodes and edges compared to classical varieties (see Table 2), although they are generated from texts of comparable length. This implies that texts in late varieties consist of fewer lemmas and on average each lemma appears more frequently (possibly creating more edges). This boosts the connectedness of the corresponding network artificially.

Irregularity can be assessed quantitatively through spectrum analysis, which consists in estimating the eigenvalues of the binary adjacency matrix of a global network. λ is an eigenvalue for this matrix on condition that there is a non-zero vector x (named eigenvector) that can satisfy the equation: $Ax = \lambda x$. The spectrum of A is the density of the set Λ including all the eigenvalues $\lambda_1 \dots \lambda_n$ and their multiplicities (the number at which an identical eigenvalue repeats): the set cardinality amounts to the number of rows/columns in the matrix. The density is a function over a continuous random variable (in this case, eigenvalues) and represents the likelihood that the variable values fall within a certain range. This likelihood is evaluated as the integral over the function values within that range.

Spectrum analysis has been proven to be useful in unraveling grammatical regularities that are independent from pure frequency by Choudhury et al. (2010). In fact, there are methods such as the Dorogovtsev-Mendes growth model (Dorogovtsev et al. 2000) to generate artificial networks that are indistinguishable from real networks created from corpora with respect to their topological properties. Crucially, however, real and artificial networks differ in their spectra. In particular, the density of the former is higher around zero. This happens because grammatical constraints make the neighborhoods of nodes in real networks more regular (see above). In other words, the rows (or equivalently columns) of elements with similar grammatical behavior are more similar. As the eigenvectors and eigenvalues can be interpreted geometrically as the direction coordinates and the factor of a transformation, respectively, then the more the factor tends to 0, the more reduced is the extent of the transformation. The lesser this extent, the more regular a matrix.

To evaluate freedom in word order, we calculated the spectra of co-occurrence and dependency networks (treated as unoriented in order that eigenvalues are real numbers). We plot the spectra for Ancient Greek in Figure 2 and the spectra for Latin in Figure 3.

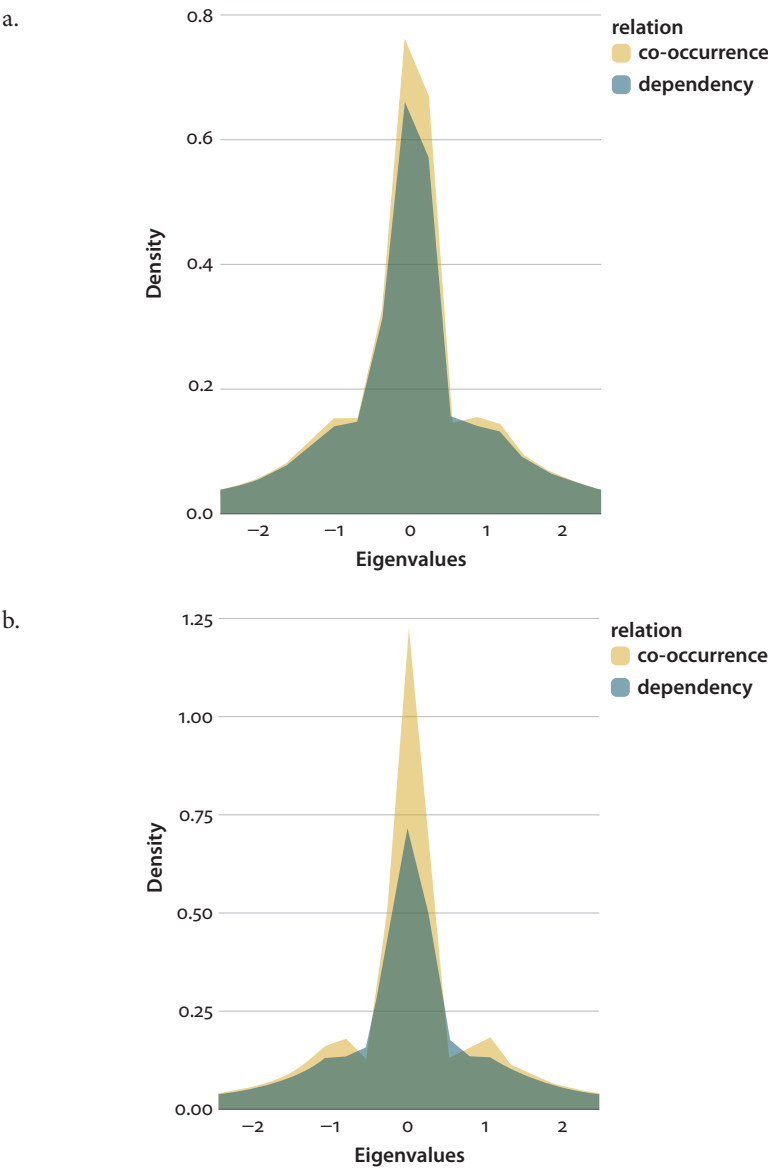


Figure 2. Spectra for co-occurrence (yellow) and dependency (green) networks for Classical (a) and Late (b) Greek

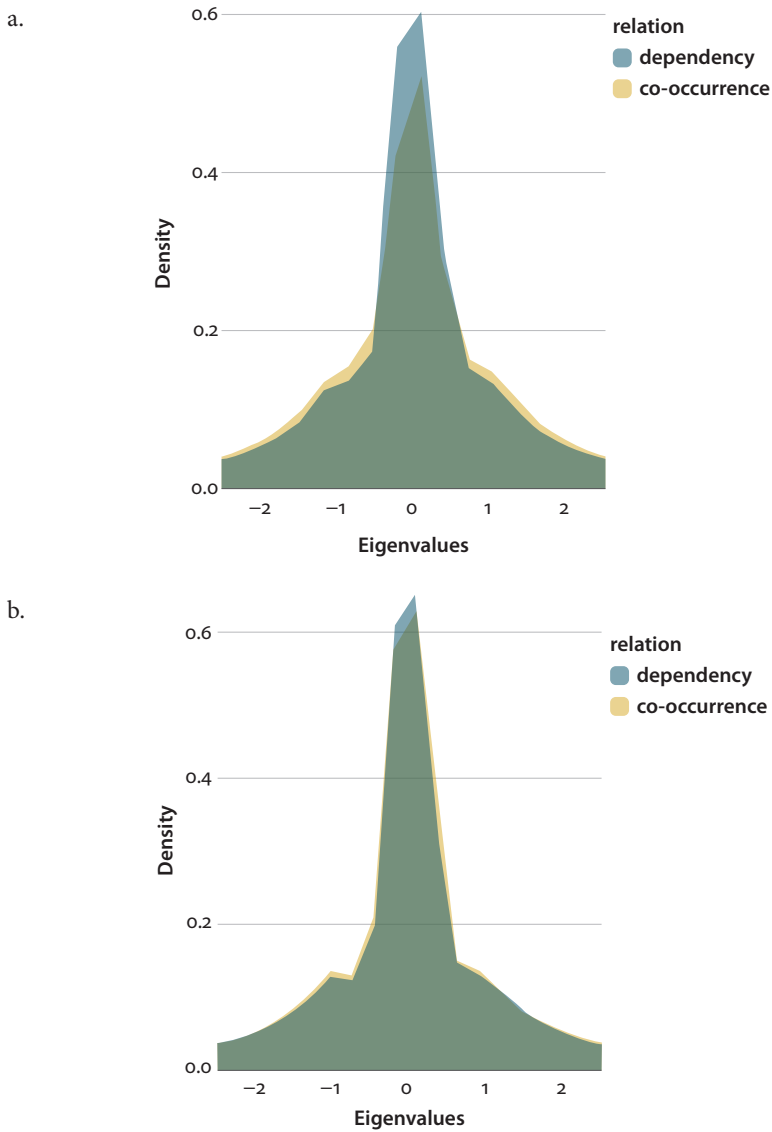


Figure 3. Spectra for co-occurrence (yellow) and dependency (green) networks for Classical (a) and Late (b) Latin

In Table 4 we compare the numerical values for the densities of $\lambda = 0$ (null factor of transformation).

Table 4. Density at $\lambda = 0$ of the co-occurrence and dependency adjacency matrixes

Variety	Co-occurrence	Dependency
Classical Greek	0.74	0.63
Late Greek	1.21	0.69
Classical Latin	0.50	0.59
Late Latin	0.60	0.63

As shown in Table 4, the density of the eigenvalues around 0 grows in late varieties for co-occurrence networks. A higher value means a higher regularity in the network; in turn, we maintain this to be a proxy for a more rigid word order. In particular, the metric value increases by 63.5% in Ancient Greek and by 20% in Latin. The soundness of our method is demonstrated by the stability of the density at 0 across languages and time for dependency networks. In fact, we expect that these syntactic relations, being universal, enforce a constant set of constraints on the word combinations.

4.2 Discontinuous NPs

We measured NP discontinuity by analyzing local networks, i.e., dependency trees. We considered two types of constituents: NPs consisting of (i) article + noun for Ancient Greek or (ii) attributive adjective + noun for both languages. These constituents were counted as discontinuous on condition that their components were separated in the linear order by at least an element that does not belong to their subtree. More formally, if a node lies between other nodes but does not share a common sub-tree with them, it creates non-projectivity and is said to be ‘in a gap’ (Marcus 1965). Note that in both the cases a head-dependent relation holds between the components of the constituent.

Mambrini & Passarotti (2013) have shown that the amount of non-projective trees in Ancient Greek and Latin is higher than in any modern Indo-European language they were able to test – based on treebank availability – on account of the occurrence of discontinuous constituents. In Ancient Greek, this phenomenon is even more relevant: the main causes listed by the authors are clitics and other P2 particles, which occupy the second position in the sentence and thus split any constituent spanning across that position. Mambrini & Passarotti (2013) found another source of non-projectivity in the displacement to the left of arguments and adjuncts of the verb either for pragmatic purposes or from a subordinate clause to the main clause. For nouns, interrogative pronouns and predicative adjectives contribute to non-projectivity the most.

In late varieties of Ancient Greek and Latin, however, continuous NPs grammaticalized. This trend has been charted quantitatively: Gulordava & Merlo (2015) estimated the percentage of adjacent heads and dependents in NPs. Diachronically they observed a sharp decrease in non-adjacent modifiers in both Ancient Greek and Latin. However, this criterion does not necessarily imply a discontinuity, since heads and dependents can be separated by words that depend on either of them and belong to the same constituent. In this work, we devise a metric limited to actual discontinuities.

We measured the absolute frequency of constituents separated by at least a ‘node in a gap’ in the four language varieties using the Tree Query extension of the TrEd Tree Editor² software.³ In order to make the treebanks readable by this software, we pre-processed them by converting them into the Prague Markup Language (PML), a data format based on XML intended for storing linguistically annotated data. In the queries, we searched nodes preceding/following a noun and following/preceding an adjective or an article with the dependency relation of attribute or determiner, respectively. We also ensured that the adjective or article depends on the noun and that the intervening nodes are not part of this subtree. As an example, consider the sentence in (12) and the ensuing tree retrieved by the query in Figure 4.

- (12) *Qui diutissime impuberes permanserunt, maximam
REL.NOM.PL long chaste:NOM.PL remain:PRET.3PL highest:ACC
inter suos ferunt laudem
among POSS.3PL.ACC.PL carry:PRS.3PL praise:ACC
“Those who have remained chaste for the longest time, receive the greatest
commendation among their people.” (Caes. G. 6.21)*

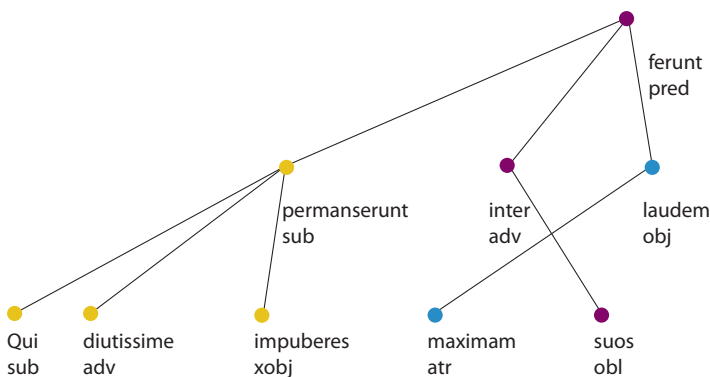


Figure 4. Discontinuous constituent in Classical Latin retrieved by the query

2. <https://ufal.mff.cuni.cz/tred/>

3. We excluded discontinuous constituents in coordinated constructions to simplify the query.

In (12), a NP *maximam ... laudem* “highest praise” (in blue in Figure 4) is separated by three words in a gap *inter suos ferunt* “receive among them” (in purple), which belong to a different subtree.

The results of the queries appear in Table 5 (the total includes right-headed pairs (AdjN) and left-headed pairs (NAdj)).

Table 5. Counts of pairs divided by at least a word in a gap (i.e., creating discontinuities) by kind of constituent

Language variety	Noun phrases (determiner)	Noun phrases (adjective)
Classical Greek	400	85 + 206 = 291
Late Greek	112	12 + 23 = 35
Classical Latin	–	221 + 58 = 279
Late Latin	–	25 + 18 = 43

Table 5 shows that the number of discontinuous constituents drops dramatically in late varieties for both kinds of nominal constituents compared to classical varieties. Discontinuous NPs with determiners drop by –72% in Ancient Greek; discontinuous NPs with adjectives drop by –87.27% in Ancient Greek and –84.58% in Latin. These results confirm the grammaticalization of constituency, which is often held to be true but seldom assessed quantitatively in the literature (e.g., Ledgeway 2011, 2012: 31–58).

As we saw in §2.2, Classical Greek NPs had definite articles while Latin did not. Table 5 indicates a major difference between the drop in frequency of split constituents depending on whether it is the article that is separated from the noun or whether it is an adjective. In the former, the frequency is less reduced than in the latter. This depends on the fact that, as remarked in §2, when an NP with a definite article occurs initially in the sentence, a P2 sentence particle is placed immediately after the article, as in (3). Indeed, a qualitative analysis shows that the nodes in a gap between the article and the noun in Ancient Greek derive from the occurrence of such particles, as shown in Table 6, in which we considered the top five nodes in a gap by the frequency (in parentheses) of their lemma.

Table 6. Ranking by frequency of words in a gap between articles and nouns in the varieties of Ancient Greek

Classical Greek	Late Greek
<i>dé</i> (254)	<i>dé</i> (177)
<i>te</i> (76)	<i>oún</i> (16)
<i>mén</i> (48)	<i>gár</i> (15)
<i>dé</i> (43)	<i>mén</i> (14)
<i>gár</i> (41)	<i>te</i> (9)

4.3 Referential null objects

To estimate the number of referential null objects, no direct source for information is available either in local or in global networks, since co-reference and implicit nodes are not present in all the treebanks. Instead, this measure was approximated by the percentage of third person personal pronouns among the objects of verbs. We assumed that, if the object is mandatory, it must surface as a pronoun in some constructions.⁴ On the other hand, when null anaphora of the object is possible, no object at all is expressed. As a consequence, the loss of null objects in late varieties is expected to be related to the skyrocketing of the rate of personal pronouns among objects. As the nominative of first and second person pronouns is especially frequent in discourse, we limit our observations to third person pronouns. Table 7 contains their count, the total object count and their ratio. Also, we counted the number of verb pairs in coordination with the following requirements: both must govern objects, the first being a noun and the second a third person personal pronoun. Although this query retrieves false positives where the two objects are not co-referent, it nevertheless indicates a clear diffusion of this pattern in late varieties.

Table 7. Counts of the objects (total and subsets)

	CG	LG	CL	LL
objects	7171	6659	6944	7675
of which 3rd person personal pronouns	295 (4.11%)	789 (11.85)	167 (2.40)	763 (9.94)
of which in coordinated constructions	5 (0.07%)	58 (0.87%)	10 (0.14%)	57 (0.74%)

The results in Table 7 strongly support our hypothesis. This change also entails a side effect in global networks. Each node in these networks is associated with a degree, which equals the number of its edges. The nodes with the highest degree are defined hubs, and their deletion alters the topology of the linguistic networks dramatically because this would make them highly disconnected. Since in late varieties personal pronouns occur more often and they collocate with an open, paradigmatically rich class like verbs, they increase the number of edges of their corresponding nodes in global networks. This is equivalent to saying that they increase their 'hubness'. Table 8 shows the top six lemmas by degree. Note that since

4. One could object that a direct object can be realized by an NP rather than a (personal) pronoun. Note however that null direct objects occur in contexts in which they are easily recoverable from the context (see Luraghi 1997, 2003 on specific conditions), and in such conditions overtly realized objects are normally weak pronouns.

these figures concern lemmas, and not specific forms, all forms of pronouns are included, most notably the nominative.

Table 8. Ranking of the most connected nodes of global networks by number of edges

Position	CG	LG	CL	LT
1	<i>ho</i>	<i>ho</i>	<i>et</i>	<i>et</i>
2	<i>kaí</i>	<i>kaí</i>	<i>sum</i>	<i>sum</i>
3	<i>dé</i>	<i>autós</i>	<i>qui</i>	<i>is</i>
4	<i>eimí</i>	<i>dé</i>	<i>que</i>	<i>in</i>
5	<i>hoûtos</i>	<i>eimí</i>	<i>is</i>	<i>autem</i>
6	<i>autós</i>	<i>egó</i>	<i>ego</i>	<i>qui</i>

Third person pronouns (*autós* for Ancient Greek, *is* for Latin) climb the ranking in the late varieties compared to the classical varieties.⁵ The evidence from both the local and global networks points towards a change in the role of third person pronouns: in particular, their rising tendency to appear as verb objects points toward the overt realization of the object even in contexts in which its referent can be recovered from discourse.

5. Discussion

In the previous sections, we drew three straightforward conclusions, which can be summarized as follows:

- Free word order. Words in late varieties show a higher regularity in the co-occurrence patterns.
- Constituents. The number of discontinuous NPs drops dramatically in late varieties.
- Null referential direct objects. Third-person personal pronoun objects increase in number inside the treebanks, and third person pronouns increase in degree inside the global networks in late varieties.

Is the variation across these parameters correlated? In order to assess this, we calculated the Pearson correlation coefficients of each possible pair of metrics. This parameter measures the linear correlation between a pair of random variables, i.e., values observed for a given parameter: in our case, the values within a single

5. Note that we are giving the lemmas of pronouns. We are, however, well aware of the fact that in no Ancient Greek variety does the nominative *autós* function as a third-person anaphoric pronoun: this function is limited to non-nominative forms.

language. Pearson correlation coefficients measure the covariance of the two random variables normalized by the product of their standard deviations. A matrix of Pearson correlation coefficients for the measures of non-configurationality features is shown in Table 9. In the bottom-left half we report the strength of the correlation (ρ). The possible values range between -1 (perfect negative correlation) and 1 (perfect positive correlation), passing by 0 (no correlation). The top-right of the table, instead, shows the statistical significance. Such confidence in the correlation is expressed by p -values, i.e., the probabilities (from impossibility 0 to certainty 1) that the correlation has emerged by chance (in other words, that the null hypothesis is true). Note that we excluded NPs with articles because part of the values are not applicable, as Latin has no articles, hence the measure cannot be computed.

Table 9. Pearson coefficients (bottom-left) and their significance expressed as p -value (top-right) of the correlations among the variables measured in both global and local networks

	Density	Discontinuous NPs w/ adjective	3rd person personal pronouns in coordination
Density	–	$p = 0.4692$	$p = 0.4884$
Discontinuous NPs w/ adjective	$\rho = -0.53$	–	$p = 0.0007$
3rd person personal pronouns in coordination	$\rho = 0.51$	$\rho = -0.99$	–

Based on the figures in Table 9, we found a correlation to be statistically significant and strong. Indeed, the count of third-person personal pronouns in coordination correlates negatively with the number of non-projective adjective-noun pairs. Since we linked the former with the absence of the null anaphora of objects, it turns out that the occurrence of null referential direct objects and the occurrence of discontinuous NPs are interdependent.

On the other hand, metrics related to word order freedom do not offer any evidence to support a correlation with the other metrics: this might be due to the independence of word order freedom from non-configurationality. In fact, word order is relatively unconstrained also in some languages that are usually taken to be configurational, such as most Romance languages, hence it is comparatively less revealing than other features. In general, the correlations show that various developments commonly considered typical of increasing configurationality besides being parallel in time are indeed part of the same ongoing change.

6. Conclusions

We have argued that correlates of non-configurationality found in Classical Greek and Latin were declining in late varieties of the two languages and that their decline can be measured by using appropriate metrics based on network analysis, both at the local level (syntactic dependency trees and word order of single sentences) and at the global level. In this way, we have identified clues to increasing configurationality, consisting in a decrease in the freedom of word order, in the almost complete disappearance of discontinuous NPs and in the increase of pronominal direct objects (mirror image of the simultaneous decrease of null direct objects). Moreover, we have found a significant and strong correlation between discontinuous NPs and null direct objects, demonstrating that these variables co-vary over time.

Nevertheless, there remain some caveats with respect to the data, metrics and variables considered. As mentioned in §3.1, diachrony is not the only variable explaining differences in the texts: style and genre can also influence the freedom of word order. In Ancient Greek, discontinuity within NPs shows a lesser decrease in cases in which definite articles occur, due to the frequency of P2 particles that were routinely placed between the article and the noun. Moreover, while we have proposed the presence of third person personal pronominal direct objects as a proxy for the absence of null direct objects in specific contexts in which their referent is recoverable from discourse, the ideal metric for null objects would also require a pragmatic level of annotation, which is only partly available for these languages currently. Finally, a larger sample in future experiments would be required to corroborate the evidence for this correlation. In particular, more languages should be taken into account, and the treebanks for Ancient Greek and Latin should be extended by manual annotation or syntactic parsing (Gulordava & Merlo 2015; Ponti & Passarotti 2016).

Acknowledgements

We would like to thank the editors of *Diachronica* and the reviewers for their invaluable suggestions, which made our work more thorough in both the theoretical assumptions and the technical implementation. Edoardo Ponti also wishes to thank Marco Passarotti for having introduced him into the fascinating field of network analysis.

References

- Austin, Peter & Joan Bresnan. 1996. Non-configurationality in Australian languages. *Natural Language and Linguistic Theory* 14, 215–268. <https://doi.org/10.1007/BF00133684>
- Baker, Mark. 2001. Configurationality and polysynthesis. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, vol. 2, 1433–1441. Berlin: Mouton de Gruyter.
- Bamman, David, Francesco Mambrini & Gregory Crane. 2009. An ownership model of annotation: The Ancient Greek dependency treebank. In Marco Passarotti, Adam Przepiórkowski, Savina Raynaud & Frank van Eynde (eds.), *Proceedings of the eighth international workshop on treebanks and linguistic theories (TLT 8)*, 5–16. Milan: EDUCatt.
- Baronchelli, Andrea, Ramon Ferrer-i-Cancho, Romualdo Pastor-Satorras, Nick Chater & Morten H. Christiansen. 2013. Networks in cognitive science. *Trends in Cognitive Sciences* 17(7), 348–360. <https://doi.org/10.1016/j.tics.2013.04.010>
- Čech, Radek, Ján Mačutek & Zdeněk Žabokrtský. 2011. The role of syntax in complex networks: Local and global importance of verbs in a syntactic dependency network. *Physica A: Statistical Mechanics and its Applications* 390(20), 3614–3623. <https://doi.org/10.1016/j.physa.2011.05.027>
- Choudhury, Munmun, Dipak Chatterjee & Animesh Mukherjee. 2010. Global topology of word co-occurrence networks: Beyond the two-regime power-law. In Aravind K. Joshi, Chu-Ren Huang & Dan Jurafsky (eds.), *Proceedings of the 23rd international conference on computational linguistics: Posters*, 162–170. Stroudsburg, PA: Association for Computational Linguistics.
- Danckaert, Lieven. 2015. Studying word order changes in Latin: Some methodological considerations. In Carlotta Viti (ed.), *Perspectives on historical syntax*, 233–250. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.169.09dan>
- Deligianni, Efrosini. 2011. Modern Greek word order in the process of syntacticization: Preliminary evidence from Late Byzantine and Early Modern Greek. In Katerina Chatzopoulou, Alexandra Ioannidou & Suwon Yoon (eds.), *Proceedings of the 9th international conference on Greek linguistics (ICGL 9)*, 440–455. Columbus: The Ohio State University.
- Devine, Andrew & Laurence Stephens. 2000. *Discontinuous syntax: Hyperbaton in Greek*. Oxford: Oxford University Press.
- Dorogovtsev, Sergey N., José Fernando F. Mendes & Alexander N. Samukhin. 2000. Structure of growing networks with preferential linking. *Physical review letters* 85, no. 21:4633. <https://doi.org/10.1103/PhysRevLett.85.4633>
- Dover, Kenneth. 1960. *Greek word order*. Cambridge: Cambridge University Press.
- Ferrer i Cancho, Ramon & Richard V. Solé. 2001. The small world of human language. *Proceedings of the Royal Society of London B: Biological Sciences* 268(1482), 2261–2265. <http://dx.doi.org/10.1098/rspb.2001.1800>
- Ferrer i Cancho, Ramon, Ricard V. Solé & Reinhard Köhler. 2004. Patterns in syntactic dependency networks. *Physical Review E* 69, no. 5:051915.
- Futrell, Richard, Kyle Mahowald & Edward Gibson. 2015. Quantifying word order freedom in dependency corpora. In Joakim Nivre & Eva Hajičová (eds.), *Proceedings of the third international conference on dependency linguistics (Depling 2015)*, 91–100. Uppsala: Uppsala University.

- Goldstein, David. 2016. *Classical Greek syntax: Wackernagel's Law in Herodotus*. Leiden: Brill.
- Gulordava, Kristina & Paola Merlo. 2015. Diachronic trends in word order freedom and dependency length in dependency-annotated corpora of Latin and Ancient Greek. In Joakim Nivre & Eva Hajičová (eds.), *Proceedings of the third international conference on dependency linguistics (Depling 2015)*, 121–130. Uppsala: Uppsala University.
- Hale, Ken. 1983. Warlpiri and the grammar of non-configurational languages. *Natural Language and Linguistic Theory* 1. 5–47.
- Haug, Dag T. T. & Marius L. Jøhndal. 2008. Creating a parallel treebank of the Old Indo-European Bible translations. In Caroline Sporleder and Kiril Ribarov (eds.), *Proceedings of the workshop on language technology for cultural heritage data (LaTeCH 2008)*, 27–34. Marrakech, Morocco.
- Hewson, John & Vit Bubenik. 2006. *From case to adposition: The development of configurational syntax in Indo-European languages*. Amsterdam: John Benjamins.
<https://doi.org/10.1075/cilt.280>
- Kapustin, Victor & Anna Jamsen. 2007. Vertex degree distribution for the graph of word co-occurrences in Russian. In Chris Biemann, Irina Matveeva, Rada Mihalcea & Dragomir Radev (eds.), *Proceedings of the second workshop on TextGraphs: Graph-based algorithms for natural language processing*, 89–97. Rochester, NY: Association for Computational Linguistics.
- Keydana, Götz & Silvia Luraghi. 2012. Definite referential null objects in Vedic Sanskrit and Ancient Greek. *Acta Linguistica Hafniensia* 44(2). 116–128.
- Kiss, E. Katalin. 1987. *Configurationality in Hungarian*. Dordrecht: Reidel.
- Ledgeway, Adam. 2011. Morphosyntactic typology and change. In Martin Maiden, John Charles Smith & Adam Ledgeway (eds.), *The Cambridge history of the Romance languages, vol. 1: Structures*, 382–471, 724–734. Cambridge: Cambridge University Press.
- Ledgeway, Adam. 2012. *From Latin to Romance*. Oxford: Oxford University Press.
- Linde, Paul. 1923. Die Stellung des Verbums in der lateinischen Prosa. *Glotta* 12. 153–178.
- Luraghi, Silvia. 1997. Omission of the direct object in Classical Latin. *Indogermanische Forschungen* 102. 239–257.
- Luraghi, Silvia. 1998. Omissione dell'oggetto diretto in frasi coordinate: Dal latino all'italiano. In Paolo Ramat & Elisa Roma (eds.), 183–196.
- Luraghi, Silvia. 2003. Definite referential null objects in Ancient Greek. *Indogermanische Forschungen* 108. 169–196.
- Luraghi, Silvia. 2010. The rise (and possible downfall) of configurationality. In Silvia Luraghi & Vit Bubenik (eds.), *Continuum companion to historical linguistics*, 212–229. London: Continuum.
- Luraghi, Silvia. 2013. Clitics. In Silvia Luraghi & Claudia Parodi (eds.), *The Bloomsbury Companion to Syntax*, 165–193. London: Bloomsbury.
- Mambrini, Francesco & Marco Passarotti. 2013. Non-projectivity in the Ancient Greek dependency treebank. In Eva Hajičová, Kim Gerdes & Leo Wanner (eds.), *Proceedings of the second international conference on dependency linguistics (Depling 2013)*, 177–186. Prague: Matfyz Press.
- Marcus, Solomon. 1965. Sur la notion de projectivité. *Mathematical Logic Quarterly* 11(2). 181–192. <https://doi.org/10.1002/malq.19650110212>
- Meillet, Antoine & Joseph Vendryes. 1924. *Traité de grammaire comparée des langues classiques*. Paris: Champion.

- Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajic, Christopher D. Manning et al. 2016. Universal dependencies v1: A multilingual treebank collection. In Nicoletta Calzolari (ed.), *Proceedings of the tenth international conference on language resources and evaluation (LREC 16)*, 1659–1666. Portorož: European Language Resources Association.
- Ponti, Edoardo M. 2016. Divergence from syntax to linear order in Ancient Greek lexical networks. In Zdravko Markov and Ingrid Russel (eds.), *Proceedings of the twenty-ninth international FLAIRS conference*, 187–193. Palo Alto: AAAI Press.
- Ponti, Edoardo M., & M. Passarotti. 2016. Differentia compositionem facit: A slower-paced and reliable parser for Latin. In *Proceedings of the tenth international conference on language resources and evaluation (LREC 16)*, 683–688. Portorož: European Language Resources Association.
- Reinöhl, Uta. 2016. *Grammaticalization and the rise of configurationality in Indo-Aryan*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198736660.001.0001>
- Revithiadou, Anthi & Vassilios Spyropoulos. 2007. *A typology of Greek clitics with special reference to their diachronic development*. Ms., University of the Aegean. <http://ling.auf.net/lingBuzz/000496> (last accessed on 11/07/2018.)
- Revithiadou, Anthi & Vassilios Spyropoulos. 2008. Greek object clitic pronouns: A typological survey of their grammatical properties. *Language Typology and Universals* 61(1). 39–53. <https://doi.org/10.1524/stuf.2008.0005>
- Rögnvaldsson, Eiríkur. 1995. Old Icelandic: A non-configurational language? *North-Western European Language Evolution* 26. 3–29. <https://doi.org/10.1075/nowele.26.01rog>
- Salvi, Giampaolo. 2004. *La formazione della struttura di frase romanza*. Tübingen: Niemeyer. <https://doi.org/10.1515/9783110945508>
- Schäufele, Steven. 1990. *Free word-order syntax: The challenge from Vedic Sanskrit to contemporary formal syntactic theory*. Urbana-Champaign: University of Illinois at Urbana-Champaign dissertation.
- Solé, Richard V., Bernat Corominas Murtra, Sergi Valverde & Luc Steels. 2010. Language networks: Their structure, function, and evolution. *Complexity* 15(6). 20–26. <https://doi.org/10.1002/cplx.20326>
- Tesnière, Lucien. 1959. *Éléments de syntaxe structurale*. Paris: Klincksieck.

Résumé

La non-configurationalité est une propriété de certaines langues, associée à l'ordre des mots libre, aux constituants discontinus et à l'anaphore nulle des arguments verbaux référentiels. Certaines métriques quantitatives, s'appuyant sur les réseaux locaux (arbre syntaxique et ordre des mots d'une seule phrase) et les réseaux globaux (qui intègrent toutes les relations locales dans un graphe partagé), peuvent révéler des corrélations entre ces attributs. Nous nous concentrons sur les variétés diachroniques du grec ancien et du latin, où la non-configurationalité s'amenuise avec le temps. Nous montrons qu'une propriété des réseaux globaux (la densité de leurs spectres autour de « valeurs propres » nulles) qui mesure la régularité dans l'ordre des mots devient plus prononcée des variétés anciennes aux variétés plus récentes. Les constituants discontinus sont mesurés en comptant les mots qui génèrent des arbres de dépendance non-projectifs : ceux-ci chutent de façon spectaculaire dans les variétés tardives. Enfin, les développements dans

l'utilisation d'anaphores nulles sont évalués à l'aide du pourcentage de pronoms personnels à la troisième personne parmi les objets verbaux. Tous ces trois attributs s'avèrent changer dans le temps en raison de la décroissance de la non-configuralité. L'évaluation de la force de leur corrélation par paire montre que les objets directs nuls et les constituants discontinus sont profondément liés.

Zusammenfassung

Nicht-Konfiguralität ist eine Eigenschaft einiger Sprachen, die mit freier Wortstellung, diskontinuierlichen Konstituenten (Nominalphrasen eingeschlossen) und Nullanapher referentieller Argumente assoziiert ist. Die auf lokalen Netzwerken (syntaktische Bäumen und Wortstellung innerhalb von Sätzen) sowie auf globalen Netzwerken (Relationen innerhalb einer ganzen Baumbank in ein gemeinsames Diagramm inkorporierend) basierende quantitative Metrik kann Korrelationen zwischen diesen Merkmalen deutlich machen. Wir richten unser Augenmerk auf diachrone Varietäten des Altgriechischen und Lateinischen, bei denen die Nicht-Konfiguralität mit der Zeit abgenommen hat, was zum überwiegend konfiguralen Charakter des modernen Griechischen und der romanischen Sprachen geführt hat. Es wird gezeigt, dass eine Eigenschaft der globalen Netzwerke (Dichte von deren Spektren um den Eigenwert Null), welche die Regelmäßigkeit in der Wortstellung misst, im Lauf der Entwicklung von den klassischen zu den späten Varietäten verstärkt wird. Diskontinuierliche Nominalphrasen werden dadurch bestimmt, indem die in Dependenzbäumen Nicht-Projektivität erzeugenden Wörter gezählt werden: Diese lassen in den späten Varietäten dramatisch nach. Schließlich wird die Entwicklung des Gebrauchs referentieller Nullobjekte durch die Bestimmung der Prozentzahl der Personalpronomina der dritten Person unter Verbalobjekten abgeschätzt. Diese drei Merkmale ändern sich mit der Zeit wegen des Verfalls der Nicht-Konfiguralität. Die Abschätzung der Stärke der paarweisen Korrelation zeigt, dass Null-Direktobjekte und diskontinuierliche Nominalphrasen tief miteinander verknüpft sind.

Authors' addresses

Edoardo Maria Ponti
St John's College
University of Cambridge
St John's Street
CAMBRIDGE, CB2 1TP
United Kingdom
ep490@cam.ac.uk

Silvia Luraghi
Università di Pavia
Dipartimento di Studi Umanistici, Sezione di Linguistica Teorica e Applicata
Strada Nuova 65
I-27100 PAVIA
silvia.luraghi@unipv.it