

2.7 Clitics¹

Silvia Luraghi

University of Pavia

1. Introduction

Clitics have given rise to an enormous number of studies. In particular, they have been (and still are) an extremely popular topic of research in various versions of GB, including Minimalism, and other formal theories, such as Optimality. Even just listing relevant contributions on clitics would take much more than the space allotted for this chapter, so the interested reader is referred to Nevis et al. (1994), and, for more recent publications, to other works cited in the references.² Here I limit myself to analyzing some aspects of two types of clitics, viz. Romance pronominal clitics and P2 clitics as instantiated in some ancient Indo-European languages, with a focus on Hittite and on Ancient Greek.

As remarked above, available treatments of clitics are mostly based on formal syntactic criteria; in addition prosodic treatments have also been worked out. In this chapter, I give a usage based account of clitics, following an approach that views prosody as reflecting discourse organization, which I outline in sec. 2. Section 3 is devoted to some aspects of Romance pronominal clitics, including clitic doubling (3.1), clitic climbing (3.2), and the rise of pronominal clitics out of former

¹ Research for this paper has been generously supported by a grant of the Alexander von Humboldt Foundation in the fall of 2011.

² An updated comprehensive bibliography of studies on clitics is not available. Note further that bibliographies of GB inspired works are biased as they virtually ignore everything which has been published outside formal theories, see for example the papers in van Reimsdijk (1999).

unaccented pronouns (3.3). In section 4 I discuss P2 clitics in two languages that display different tendencies regarding their placement: Hittite, in which P2 positioning of clitics is almost exceptionless (4.1), and Classical Greek, which presents more varied patterns, highly conditioned by discourse structure (4.2). The remainder of this section is devoted to various issues related to clitics, including the difference between ‘simple’ and ‘special’ clitics (1.1), different placement possibilities (1.2), possible types of host (1.3) and order of clitics in clusters (1.4).

1.1. Simple vs. special clitics

There is little consensus on what clitics are, how they can be classified into types, and even on their existence. According to Joseph (2002), for example, all items identified as clitics in current research on Modern Greek (which are largely similar to Romance clitics discussed in sec. 3) are best understood if considered either words or affixes. In this chapter, I follow the more widespread opinion that clitics are some sort of in-between items, which share some features of bound and some of free morphemes.³

Indeed, clitics are items that cannot stand independently of other (‘real’) words in a sentence as they most often do not bear an accent. Prosodically, they are ‘deficient’, and constitute a prosodic unit, or clitic group with their host.⁴ In the meantime, clitics can display a morphological behavior

³ As I argue in sec. 2, in a usage based perspective it makes little sense to try to define words, affixes, even constructions and, of course, clitics, as sharply separate categories, since linguistic units are better understood as a result of chunking processes based on frequency of collocations; see below, sec. 2.

⁴ Clitics can be shown not to constitute phonological words with their hosts as they normally do not trigger internal phonological processes triggered by bound morphemes (see among others Nespor, Vogel 2007 and Russi 2008: 211-217 on Italian enclitics). The complex constituted by clitics with

typical of words, as e.g. be inflected in relevant grammatical categories: this is the case of pronouns or auxiliaries, two lexical classes that often belong in the inventory of a language's clitics. Thus, one way of characterizing clitics is by saying that they are not phonological words, but that morphologically they must be considered words in their own right.

Since Zwicky (1977) it is common to distinguish between 'simple' and 'special' clitics. According to Zwicky's definition, simple clitics are unaccented variants of 'real' words that lose their lexical accents in specific conditions, as for example the unaccented form [əm] of English third person non-subject pronoun *him*. On the contrary, special clitics are not simply unaccented variants of lexically accented words: they are lexically unaccented items, which have a distribution of their own, and, crucially, peculiar placement rules which cross-linguistically appear to be limited to a small number of options. Romance pronominal clitics and second position, or P2 clitics best illustrate special clitics in this sense, partly on account of the fact that Zwicky's definition is mostly worked out in order to accommodate specificities of Romance clitics; as we will see in sec. 4.2, some P2 clitics have a somewhat more controversial status as they are lexically accented though confined to second position or at least ruled out from first position.

1.2. Two options for clitic placement

Cross-linguistically, clitic placement is determined by two possible tendencies:

- (a) clitics attach phonologically to specific constituents (constituent based, or constituent hosted clitics);

their host is often called clitic group (e.g. Nespor, Vogel 2007). In his description of Classical Greek clitics, discussed in sec. 4.2, Goldstein (2010) uses the term 'prosodic word' in a similar sense.

(b) clitics attach to a specific position in a sentence, typically P2.⁵

Klavans (1985) considers cliticization as a primarily syntactic phenomenon, with phonology only playing a role as a last factor. She considers the domain of cliticization to be determined syntactically, and draws a distinction between the structural host of a clitic, that is, the constituent to which the clitic belongs syntactically, and its phonological host, that is, the constituent to which the clitic attaches phonologically. In Klavan's terms, type (a) clitics attach to the same constituent both phonologically and syntactically, while P2 clitics may attach to two different hosts (i.e. a phonological host and a distinct structural host at the same time).

Among clitics mentioned in sec. 1, Romance pronouns belong to type (a) as they always attach to the verb (parameters of possible variation include enclisis vs. proclisis and the position relative to auxiliaries and modal verbs). Other constituent based clitics include focalizers, which typically attach to the word or constituent that they have in their scope, as Ancient Greek *ge* in example (1):

(1) *sù tò prosrēthēsómenon orthôs ámeinon hēdonēs _____ ge*

2SG.NOM ART.ACC say:FUT.PTCP.MID.ACC rightly better:ACC pleasure:GEN.F PTC

agathòn eînai nóûn

good:ACC be:PRS.INF mind:ACC

“You assert that the good which is rightly to be called better even than pleasure is mind.”

(Pl. *Phlb.* 19d).

NP coordinators can be clitic and attach to one of the coordinated items, as the conjunction *-que* in Latin (see Wanner 1987 on this and other constituent based clitics in Latin):

⁵ As I argue in sec. 4.1 and 5, a case can possibly be made for P1 as another preferred position for clitics in addition or alternative to P2.

(2) *quae in amicitiam populi Romani dicionem=que essent*

REL.NOM.PL in alliance:ACC people:DAT Roman:DAT subjecthood:ACC and be:SUBJ.IMPF.3PL

‘which are allies and subjects of the Roman people’ (Cic. *div. in Caec.* 66).

Possessives also tend to attach to the NP that constitutes their head, as shown in Hittite:

(3) *nu= us appa ishi=_____ssi pennai*

CONN 3PL.ACC back master.D/L 3SG.POSS.D/L drive:PRS.3SG

“(S)he takes them (= the oxen) back to their (sg.) owner.” *HG* § 79 (=Friedrich, 1959).

(Note that such Hittite possessives are adjectives, and agree with their head noun in gender, number and case. Possessive clitics realized as genitive or dative forms of pronouns in languages with P2 clitics may display a preference toward NP cliticization as in Serbo-Croatian, see Mišenska Tomić 1996, but they can also be placed in P2 as in Hittite, see further sec. 4.1).

Type (b) clitics are typically enclitic and are placed in P2: apparently, no other dedicated position for clitics is available, even though, as I argue in sec. 5, a case can perhaps be made for first position, or P1. For such clitics, the sentence is the cliticization domain in terms of Anderson (1993). Anderson further defines two other parameters of clitic placement, dominance and precedence. Dominance specifies whether the clitic attaches to the initial or final hedge of the domain, while precedence specifies whether it precedes or follows the host specified by dominance, that is, whether it is proclitic or enclitic (this parameter corresponds to phonological liaison in Klavans 1985). It follows that P2 clitics all attach to the initial hedge of their domain: they follow the item placed in P1, thus taking the leftmost position for enclitics which, by their nature, cannot stand in first position.

Focusing on clitics that can occur as (a) or type (b) cross-linguistically, as for example pronouns,

it can further be remarked that the two placement strategies can be explained as owing to competing motivations. In the first case, clitics lean phonologically on the constituent to which they belong in terms of construction and argument structure, which helps understand their function (grammatical relation or semantic role) in a sentence. In the second case, clitics tend to occur early in a sentence because of their function in discourse. The two types of motivation are in conflict with one another and drive clitics to attach to different hosts; such competing motivations can also be the cause of change in clitic placement, as discussed in sec. 4.2.

1.3. Possible hosts for P2 clitics

What counts as first position, and consequently how second position must be understood, varies. Wackernagel's Law is usually stated in terms of words: it is normally said that P2 clitics attach to the first accented word in a sentence. However, this statement needs further refinements. In the first place, languages vary as to whether they admit discontinuous constituents, and accordingly as to whether clitics can be placed after the first word of a more complex constituent. Diesing (2010) divides languages into three types, depending on whether P2 is the position after the first accented word as in Hittite and Old Bulgarian, after the first constituent, as in Czech, Slovenian, Malagasy, or whether it can be both, as in Serbo-Croatian, Ngiyambaa, Warlpiri (see further Marušič 2008 on Slovenian, and Boškovič 2001 on Serbo-Croatian). Even languages in which discontinuous constituents occur may display a tendency for certain constituents to build one single unit for the purpose of clitic placement. One such language is Classical Greek, in which determiners tend not to be separated from their noun by clitics, as I show in sec. 4.2.

In addition, evidence from various languages and different language families indicate that P2 clitics can be hosted by words (normally sentence connectives) which are elsewhere considered proclitic. One such example is provided by Hittite, which I discuss in sec. 4.1; Halpern (1995: 215) discusses similar evidence from Bulgarian. It appears that in such cases the initial proclitic can

receive an accent, or that the complex cluster proclitic+enclitic(s) can count as a clitic group. A possible explanation is that, while P2 enclitics are mostly special clitics, initial proclitics in these clusters are simple clitics, which can have an accented variant depending on the prosodic surrounding. In any case, this shows that placement of clitics in P2 cannot simply be accounted for by saying that clitics follow the first accentually strong position in a sentence: this can be the case when clitics follow an autonomous accented word, but when they contribute prosodic material for an otherwise unaccented item one has the feeling that the whole initial unit has a prosodic status of its own. I elaborate on this issue further in sec. 4.1.2. and 5.

1.4. *Order of clitics in clusters*

Clitics follow rather strict placement rules in clusters. Though such rules are language specific, they often display general tendencies, partly motivated by the scope of co-occurring clitics, partly by other factors.

Pronominal clitics typically appear in fixed slots, connected with their morphology or with person. Notoriously, there are restrictions on the co-occurrence of certain pronominal forms in clusters, typically involving second person singular accusative and third person singular dative. Compare the following Italian examples (see Russi 2008 for more examples and discussion):

(4) *Lo presento a te. / Te lo presento.*

“I introduce him to you.”

(5) *Ti presento a lui. / *Gli ti (*ti gli) presento.*

“I introduce you to him.”

Similar restrictions are reported from Modern Greek and Bulgarian, as well as from various non-Indo-European languages. They are reviewed in Haspelmath (2004b), where it is suggested that the

motivation is ultimately to be found in discourse based frequency effects.

When clusters include both constituent based and P2 clitics, the former are expected to precede the latter, and they often do, as for example in Hittite (Luraghi 1990a, b), Sanskrit (Hale 1987), and other Indo-European languages. However, this is not always the case as Pashto is reported to have P2 clitics inserted between a possible initial verb and verb-hosted deictic clitics (Tegey 1977). Since Pashto also has P2 clitics inserted between verb bases and ancient preverbs no longer analyzable as such, it seems that Wackernagel's Law can split up prosodic and even phonological words in this language (see Tegey 1977 for a description and Hock 1996 for discussion; a recent account in the framework of LFG is given in Bögel 2010).

In chains of P2 clitics it is often the case that discourse particles precede pronominal clitics: arguably, this happens because discourse particles have the whole sentence in their scope, and attach to the left sentence boundary not on account of being P2 clitics but rather because that is their syntactic domain of cliticization (structural host in terms of Klavans 1985). Again, Hittite, Ancient Greek and Sanskrit provide evidence for this tendency (see sec. 4), but other orders are also attested, as in Pashto (Tegey 1977) and Serbo-Croatian (Radanović-Kocić 1996). Given the limits of this chapter, it is impossible to even attempt here to exhaustively survey all possible orders found in clitic chains. Knowing the origin of certain clitics does not seem sufficient to shed light on specificities of their position in clusters, as argued in Steele (1977); however, diachronic developments can indicate different degrees of grammaticalization reflected in stricter or less strict cliticization. As discussed in 4.1, some Hittite particles have enclitic and apparently free standing variants, which possibly reflect a later development toward cliticization with respect to other items which only have enclitic forms.

2. Words, affixes, simple clitics, special clitics: a discourse perspective

Before moving on to the analysis of specific clitics, the issue concerning the categorial status of these items needs to be addressed. As remarked above, clitics typically bear no lexical accent in spite of displaying various morphosyntactic features typical of free rather than bound morphemes. This characteristic is puzzling only inasmuch as one's theoretical orientation forces one to work with discrete, rather than fuzzy categories, and if one fails to consider the type of information conveyed by clitics of different sorts, together with their role in discourse. Clitics are mostly items that convey grammatical, rather than lexical information, such as auxiliaries or modal particles; as repeatedly remarked, pronouns also typically have clitic forms or variants. Thus, clitics have a status which is close to that of bound morphemes: in a language like English, for example, auxiliaries have the same grammatical function as inflectional affixes in morphologically richer languages. In much the same way, clitic pronouns can be compared, and in fact often are, to cross-reference bound morphology in incorporating languages (Anderson 2003). Still, with respect to such bound morphemes, clitics display some bigger degree of freedom, and their peculiar placement rules are not such that they can easily be equated to bound morphology.

As is often the case, looking at things in a diachronic perspective can provide enlightening insights. In such perspective, clitics can be shown to be located on a grammaticalization cline that leads from free forms to bound morphology. This is the case if we consider the development of pronominal clitics in various Romance varieties (see below sec. 3.3). Clitics that express grammatical categories are likewise undergoing a grammaticalization process that might lead them to eventually become bound morphemes: as well known, verbal morphology can be shown to have originated in this way in some well documented instances, such as in the case of the Romance future tense (see Hopper, Traugott 2003). As noted by Bybee:

“The variation and gradience in the category of ‘grammatical morpheme’ is a direct result of the processes of change that affect morphemes and shape their properties of form and meaning. Lexical morphemes can become grammatical morphemes in the process of grammaticalization (as when the lexical morpheme *go* becomes part

of the future construction *be going to*), and in this process gradually become more dependent upon and eventually fused with surrounding material” (2010: 4)

Thus, wordhood is better understood as a gradience phenomenon, while drawing a sharp distinction between categories looks pointless, as shown by the following quote from Zwicky:

“This accentual test [i.e. being unaccented, SL] is probably the most popular rule-of-thumb for distinguishing clitics from independent words, but it is most unreliable; it should never, I think, be used as the sole (or even major) criterion for a classification, though it can support a classification established on other criteria. The test has two problems, one minor and one major. ... The major problem is that many clearly independent words—e.g. English prepositions, determiners, and auxiliary verbs of English—normally occur without phrasal accent” (Zwicky 1985: 287).

In spite of Zwicky’s flat statement, it is not clear at all why, for example, English prepositions are independent words more than other simple (or even special) clitics: on what criterion is wordhood established? This as well as other similar arguments in Zwicky’s paper gives a strong impression of circularity.

Clitics have in common a low degree of communicative saliency and a high degree of expectedness: for example, clitic pronouns do not introduce new information into the discourse as they rather serve the purpose of tracking referents which have already been introduced; they are highly expected because they are by the most part verb arguments. In a similar way, auxiliaries do not contribute semantic content to the predicate: rather, a typical initial step in the change from main verb to auxiliary is dessemanticization. They specify some sort of grammatical category of the main verb, and are expected because they typically occur with a verb form that needs to be specified, as a participle or an infinitive in English.

In the flow of discourse, items receive different types of accentuation depending on their communicative saliency. This simple fact, which in various forms has been re-discovered over and

over again by linguists over decades and even centuries, leads to a possible re-visitation in discourse terms of the distinction between simple and special clitics. As remarked above, following Zwicky, simple clitics are unaccented variants of lexically accented words. Several other authors have pointed out that some of Zwicky's simple clitics have a different distribution with respect to their accented counterparts, as in the case of English auxiliaries discussed in Anderson (2003), and thus present some characteristics of special clitics. This seems to suggest that the distinction between simple and special clitics is not sharp: as pointed out above, there is a gradience between lexically accented words that get unaccented or are weakly accented in discourse, and lexically unaccented items such as special clitics. This should come as no surprise: in general, linguistic phenomena have a gradient nature, something which is especially clear when one takes a discourse perspective.

Following this approach, one can find a unified explanation for other unaccented or weakly accented items as well. Indeed, lexical items that are not typical clitics also tend to lose their accent in discourse: for example, the finite verb is usually unaccented in English (or weakly accented in comparison to its arguments), unless it is focused. Compare:

(6) *Jóhn eats cákes.*

(7) *Jóhn éats cákes* (rather than baking them, for example).

Indeed, as is known at least since Wackernagel's classic paper on P2 clitics (Wackernagel 1892), the Sanskrit finite verb in main clauses displayed some features of clitics without being a clitic *stricto sensu*. In sum, words are given a different prosodic relevance depending on their discourse saliency; some items that very frequently happen to have low saliency, such as pronouns, may develop special variants that occur when they do not need to be given accentual relevance. Eventually, they may also become bound morphemes and develop into agreement markers. As noted above, the distinction between simple and special clitics is gradient: while Romance pronouns are certainly special clitics, English auxiliaries mostly conform to the definition of simple clitics, in

spite of having some features of special clitics. As we will see, among P2 clitics, too, there are items that conform more or less closely to Zwicky's definition of special clitics.

In the perspective adopted here, accentuation and deaccentuation are emergent phenomena, which happen in actual events of communication. Items which have lexically accented and lexically unaccented variants are those for which both variants are needed in discourse (such as pronouns); some items only have unaccented forms because they never need to be given prominence (some modal particles, for example), other only have accented forms, because they are virtually always salient enough to be accented (for example personal names). Following such an approach, even the issue regarding the status of clitics as free or bound morphemes loses relevance: clitics are items that frequently co-occur with other items or in special positions; as such they are 'chunked' together with their hosts, and may eventually end up being reinterpreted as bound affixes as a consequence of chunking. Chunking is a fundamental principle of language production and processing. As pointed out by Bybee (2010: 33-34), "[w]hen two or more words are often used together, they also develop a sequential relation ... The principal experience that triggers chunking is repetition." Thus, frequent collocations lead to entrenchment of constructions and low saliency items tend to acquire a bounded status in constructions, a first hint of which is their close prosodic connection with more salient items.

Prosody is an important reflex of discourse organization not only in terms of alternation between what is accented and what is not, but also because prosodic units reflect information structure in sentences. In this respect, intonational phrases are especially relevant because they represent units of information; extraposed topics or focused constituents can best be indicated in discourse in terms of intonational phrases. As Nespors and Vogel (2007: 187-188) remark, intonational phrases are highly variable. In fact, a sentence is normally a single intonational phrase, but, depending on the style, and especially on the discourse relevance of specific constituents, it can be split into more intonational phrases, as in (example from Nespors, Vogel 2007: 194):

(8)

- (a) [_I*My friend's baby hamster looks for food in the corners of its cage*]_I
- (b) [_I*My friend's baby hamster*]_I [_I*looks for food in the corners of its cage*]_I
- (c) [_I*My friend's baby hamster*]_I [_I*looks for food*]_I [_I*in the corners of its cage*]_I

As we will see in 4.2, this and other prosodic notions can partly account for the placement of P2 clitics in Classical Greek.

2. Between clitics and affixes: Romance clitic pronouns

Romance pronominal clitics are perhaps the best known and thoroughly described group of special clitics, and research regarding their status and behavior in different Romance varieties, including different historical stages, abound. They are typically hosted by the verb, and, in the modern Romance languages, are most often proclitic; they may be enclitic with some non-finite verb forms and imperatives. Clitic pronouns are not only direct and indirect objects, but can also indicate various local relations and beneficiary (*dativus ethicus*); however, their interpretation is always dependent on the verbal predicate: in this sense, the verb constitutes their domain.⁶ Some examples

⁶ Adverbial clitics refer to spatial notions that arguably depend on the verb for interpretation, as shown in (9), in which *ne* depends on the verb *venire* 'come', which requires a source expression, and (10), in which *ci* depends on *andare* 'go', which requires a goal. In addition, various Romance languages such as Italian and French feature partitive clitics, which occur in argument position or modifying an argument, such as *ne* in (11). The only type of constituent which is apparently not part of the VP is the beneficiary clitic (including the *dativus ethicus* type), as *ci* in (11). This type of

from Italian are below:

(9) *Non me ne è venuto niente di buono.*

not me(cl) from.it(cl) is come nothing of good

“Nothing good has come out of this for me.”

(10) *Paolo va a scuola tutti i giorni, Maria ci va quando le pare.*

Paul goes to school all the days Mary there(cl) goes when her(cl) seems

“Paul goes to school every day, mary goes when she likes to.”

(11) *Mia sorella ci ha cucinato una torta, sua cugina ne ha cucinate due.*

my sister us(cl) has cooked a cake her cousin part(cl) has cooked two

“My sister backed a cake for us, her cousin cooked two.”

3.1. Clitic doubling

Some Romance varieties, most notably French, also feature subject clitics. In standard French, such clitics are in complementary distribution with full NPs or accented pronominals, as shown in the following example:

(12) *Il m' a donné un livre. / Jean m' a donnè un livre.*

he(cl) me(cl) has given a book / John me(cl) has given a book

“He gave me a book. / John gave me a book.”

beneficiary is always possible, and can occur with any type of verb. In some formally oriented descriptions, it is suggested that its occurrence opens an extra slot in the subcategorization frame of the verb.

In some other varieties, including the Northern Italian Gallo-Italic vernaculars, subject clitics have gone a step further in grammaticalization as they are no longer omissible and co-occur with possible subject NPs, as for example in Piemontese:

- (13) *It vade a ca. / Ti 't vade a ca / *vade a ca.*
you(cl) you-go to home / you you(cl) you-go to home / you-go to home
“You are going home.”

Clitic doubling provides evidence for increasing grammaticalization of Romance pronominal clitics, and constitutes an argument in favor of an analysis that considers them close to bound morphemes. Note however that Romance clitics have a bigger autonomy than bound morphemes even in varieties that display clitic doubling: for example, the Northern Italian vernaculars all allow subject clitic drop at least to some extent.⁷ This ongoing tendency is possibly a consequence of the increasing role of standard Italian as a prestige variety even among speakers of vernaculars; it seems to indicate that such clitics, in spite of the high degree of boundedness reached, never really achieved the status of bound morphemes, and are still analyzed by speakers as featuring a lower degree of boundedness than, for example, verbal endings.⁸

Clitic doubling is not limited to subjects: in Spanish, for example, recipients are most often

⁷ As indicated in the example, subject drop is not allowed in standard Piemontese for second person singular; note that this does not depend on the need to disambiguate the subject as second person singular is unambiguously indicated by the verbal ending. On subject clitics in Northern Italian vernaculars see Renzi, Vanelli (1983).

⁸ In GB inspired description such subject clitics are considered inflectional affixes, and the Northern Italian varieties are considered pro-drop, contrary to standard French. See for example Brandi, Cordin (1989).

doubled by clitics, even though this is not obligatory, as shown in:⁹

- (14) *Le doy el libro a Juan / Doy el libro a Juan*
him(cl) I-give the book to John I-give the book to John
“I give John the book.”

A construction that triggers clitic doubling in many Modern Romance languages is left dislocation of topicalized constituents, as in:¹⁰

- (15) *Il giornale l' ho letto dopo cena.*
the newspaper it(cl) I-have read after dinner
“I read the newspaper after dinner.”

In Modern Italian, clitic doubling is obligatory in such sentences, which are uttered without intonational breaks between the left dislocated constituents and the rest of the sentence (i.e. they normally constitute a unique intonational phrase). Such occurrences show that clitic doubling serves the purpose of indicating an initial direct object which is not focused; it can be considered a way to deal with free word order by means of a head marking strategy, rather than by means of the dependent marking strategy (case) typical of Latin.

⁹ See Belloro (2007) for a recent discourse based analysis of clitic doubling in Spanish, which the author views as connected with cognitive accessibility of referents, and for further reference on the topic.

¹⁰ Clitic doubling is a common phenomenon in other languages with pronominal clitics, too, and occurs for example in Slovenian and Serbo-Croatian, see Marušič (2008), as well as in Modern Greek, on which see below, sec. 4.2.

3.2. *Clitic climbing*

Another remarkable phenomenon connected with Romance clitics is climbing. In cases in which an infinitive is governed by a modal or quasi-modal verb, clitics syntactically hosted by the infinitive should attach to it as enclitics, but often they are realized as proclitic, and precede the governing verb, as in:

- (16) *Incominciava a dirmelo. / Me lo incominciava a dire.*
s/he-started to tell me(cl) it(cl) / me(cl) it(cl) s/he-started to tell
“S/he started telling me that.”

Following the usage based approach adopted here clitic climbing can have two explanations: in the first place, it indicates an at least incipient degree of auxiliarization on the side of the governing verb. While the sentence above it not such that one would easily find it in the written language, with modal verbs, which are much closer to auxiliaries, climbing is the norm. In the following example the verb *sapere*, ‘know’ has the modal meaning ‘be able to’, and favors climbing in all registers:

- (17) *Non te lo saprei dire. / Non saprei dirtelo.*
Not you(cl) it(cl) I-would.know say / not I-would.know tell you(cl) it(cl)
“I wouldn’t be able to tell you.”

The second reason why clitics should tend to climb is connected with their information status: as already noted, clitic pronouns convey low saliency information, already shared by the speaker and the hearer. For this reason, they tend to appear early in the sentence, or at least at the left hedge of their phonological host. Not surprisingly, this is also the motivation for P2 pronominal clitics to

occur where they do, as I argue in sec. 4. In much the same way as subject clitics drop in Northern Italian vernaculars, clitic climbing also indicates a lower degree of coalescence of clitics with their hosts than the degree of coalescence displayed by bound affixes with respect to their lexical base. Again, clitics retain their in-between status and defy attempts to rigid categorization.¹¹

3.3. *From Latin weak pronouns to Romance clitics*

In Latin, pronouns that did not convey new or salient information were most likely simple clitics, just as English unaccented variants of personal pronouns. As Latin is basically a free word order language, it is not as easy as for English to compare their distribution with that of accented pronouns. Diachronic studies on the development of Romance clitics from Latin unaccented or weakly accented pronouns mostly focus on first and second person forms (see for example Wanner 1987 on Romance languages in general, and Salvi 2004 on Italian); for this reason, in what follows I describe the evolution of third person pronouns.

Even though sentence initial position was normally reserved to topicalized or contrasted constituents in Latin (see Spevak 2010), it is not completely clear that unaccented or weakly accented versions of pronouns, especially in the case of third person, could not stand in initial position. Consider the following examples from Sallust:

- (18) *Terrebat eum natura mortalium avida imperi*
 frighten:IMPF.3SG 3SG.ACC nature:NOM man:GEN.PL eager:NOM power:GEN

“(Though he was at first gratified with these circumstances, **Micipsa**, considering that the merit of Jugurtha would be an honor to his kingdom, yet, when he reflected that the youth

¹¹ Clitic climbing in Italian is a comparatively neglected phenomenon. A pioneering study is Berretta (1985), which however is not corpus based, and relies on few occasionally collected examples.

was daily increasing in popularity, while he himself was advanced in age, and his children but young, he was extremely disturbed at the state of things, and revolved it frequently in his mind.) The very nature of man, ambitious of power, gave **him** reason for apprehension.”

Sall. *Iug.* 6.3;

(19) *Eos omnis praeter Turpiliium inter epulas obtruncant*

3PL.ACC all.ACC except Turpilius:ACC among banquet:ACC.PL slaughter:PRS.3PL

“(When the time came, they invited **the centurions and military tribunes**, with Titus

Turpilius Silanus, the governor of the town, to their several houses.) They butchered **them**

all, except Turpilius, while they were dining.” Sall. *Iug.* 66.3.

Indeed there appears to be no difference in the information status of *eum* in (17) and *eos* in (18), and if one were to translate these passages into a Romance language, one would use clitics in both cases:

(18’) *Lo spaventava la natura umana stessa, avida di potere.*

him(cl) frightened the nature human itself eager of power

(19’) *Li massacrarono tutti, tranne Turpilio, mentre cenevano.*

them(cl) butchered all except Turpilius at+the banquet

In the light of the above examples, it is hard to formulate rules regarding the placement of unaccented pronouns in Latin, even though they did indeed most frequently appear close to the left hedge of the sentence. Very often they turned up occurring after the first word (often a conjunction) as in (18), however they could as well appear in initial position as in (19). Clearly, this tendency reflects the information status of weak pronouns: such pronouns convey shared information, which connects a sentence to the preceding discourse. As a consequence, they tend to occur early in order to fulfill their connecting function. In addition, there is a tendency for Latin unstressed pronouns,

when they do not follow a connective, to follow a discourse prominent constituent, as pointed out in Adams (1994), who accounts for the placement of such pronouns in terms of sentence focus (and indeed this could be the case for *eum* following *terrebat* in 18).

Beside overt pronominals, Latin also had null referential anaphoras of non-subject constituents (Luraghi 1997). They occurred under various circumstances, partly discourse conditioned, partly connected to specific constructions. One such construction is sentence coordination, in which null anaphora of the direct object is obligatory in the second clause if the direct objects is the same as in the first, as shown in:

- (20) *natam sibi ex Poppaea filiam Nero ultra mortale gaudium*
 born:ACC REFL.DAT from Poppaea:ABL daughter:ACC Nero:NOM beyond human:ACC joy:ACC
accepit appellavit= que Augustam
 receive:PF.3SG name:PF.3SG and Augusta:ACC
 “Nero welcomed **a daughter** from Poppaea with more than human joy and named **her** Augusta.” Tac. *Ann.* 15.23.

With third person objects, the few occurrences of an overt anaphora (a form of the third person pronoun *is*) in the second conjunct of a coordination construction present special pragmatic conditions; most often, the anaphora is placed in initial position and hosts the clitic conjunction *-que*, a hint to the fact that it is accented (see Luraghi 1997 for examples and discussion). Thus, Romance clitics did not only substitute Latin unaccented variants of pronouns, but also null non-subject pronouns.

To sum up, Latin had nothing comparable to Romance clitics, not only on account of the fact that unaccented pronouns in Latin were not special clitics, but also because their position was apparently unconstrained: they display a tendency to occur early in the sentence on account of low information load, and most often they did indeed occur after the first word, but were not limited to this position;

in addition, Latin could featured zero forms which were later replaced by overt clitic pronominals.

When we turn to the Old Romance varieties, the scene changes completely: clitics are already there, and their position is highly constrained as they have to be adjacent to the verb.¹² By default, clitics of the Old Romance languages are proclitic, but they follow the verb and become enclitic if the verb stands in sentence initial position. This placement rule is known as Tobler Mussafia's Law, and is often regarded as a heritage of Wackernagel's Law (see for example Salvi 2004). However, as we have seen above, there is evidence that Latin weak pronouns were not banned from sentence initial position.¹³ A possible scenario then would be that on their way to become special clitics, such weak forms of pronouns lost their relative placement freedom and became enclitics. As the position after the first word in a sentence was already the most frequent one for weak pronouns, they became P2 clitics. I am not going to elaborate further on this issue, but it still needs to be stressed that this possible development has to be combined with increasing tendency of clitic pronouns to become verb based.¹⁴

¹² Some example of clitics separated from the verb in Old Spanish are discussed in Fontana (1993, 1996); similar occurrences can also be found in other Old Romance varieties, see Benincà, Cinque (1983) on Old Italian. Remarkably, all such occurrence feature clitics that, while they may lean back enclitically on the first word in the sentence, precede the verb rather than follow it. I come back to this issue in sec. 5.

¹³ Some traces of Wackernagel's Law as inherited from PIE are available from Early Latin, see Wanner (1987), but it does not seem likely, in the light of the evidence provided by classical texts, that possible P2 enclisis of weak pronouns in Late Latin could be directly connected with them, and that Wackernagel's Law as inherited from PIE could have feeded the Tobler Mussafia's Law.

¹⁴ Theories on how exactly the Tobler Mussafia's Law originated are more numerous than what I can summarize here: according to Fontana (1996), for example, the reason why clitics moved to the

Salvi (2004: 178) also suggests that the disappearance of null direct objects was one of the triggers of the development of clitics. This may well be the case: null objects were one of the features of non-configurationality displayed by Latin and by the ancient Indo-European languages in general, as shown in Luraghi (2010a). In Old Italian, null objects could still occur in cases of VP coordination and even sentence coordination (Luraghi 1998, Egerland 2002), while in Modern Italian object clitics can be omitted only in case of verb coordination and are felt as marginally acceptable by some speakers. As already remarked in sec. 3.2, clitics provide a head marking strategy to signal grammatical relations, which partly substitutes the dependent marking strategy of Latin. This has been argued by several authors especially with regard to frequent clitic doubling in spoken French, see Bossong (1998) and Luraghi (2010a) for discussion.

4. Clitic and postpositives in P2

In the course of the 19th century, Indo-European scholarship repeatedly pointed out that the ancient Indo-European languages featured an array of unaccented or weakly accented items that consistently occurred in second position in sentences. It was Jacob Wackernagel who best described what is now known as Wackernagel's law in his 1892 article,¹⁵ in which he also offered a possible

left sentence boundary is ultimately to be seen in the V2 phenomenon, which is thought to have affected the Old Romance languages.

¹⁵ Goldstein (2010) gives an accurate summary of 19th century scholarship regarding the position of clitics, and argues that the clearest statement of Wackernagel's law is indeed to be found in works by Delbrück (see e.g. Delbrück 1878).

explanation for the Germanic V2 phenomenon.¹⁶

Wackernagel mostly relied on data from Homeric Greek and Vedic Sanskrit, two languages in which the tendency for clitics to be placed in P2 is quite well attested, but that also display a number of other patterns, as I will discuss below. A couple of decades later, when Hittite was deciphered in 1916, new and more compelling evidence appeared. Indeed, Hittite is cross-linguistically one of the best examples of strict obedience to Wackernagel's Law as it features long clitic chains in P2 and shows almost no exceptions. For this reason I start by illustrating the Hittite data.

4.1. *Clitic chains in Hittite*

Let us start by examining the following passage from a Hittite text:

(21)

(a) *nu= mu memir*

CONN 1SG.OBL speak: PRET.3PL

(b) *paiueni= war=an= kan kuennumeni*

go: PRS.1PL PTC 3SG.ACC PTC kill: PRS.1PL

(c) *nu= wa=tta SAG.DU-an utumeni*

CONN PTC 2SG.OBL head:ACC bring: PRS.1PL

¹⁶ Wackernagel observed that the Sanskrit verb is unaccented or weakly accented in main clauses, and suggested that the V2 phenomenon in Germanic could be a reflex of such property, ultimately going back to PIE. I cannot pursue this issue further here; the idea that P2 clitics and V2 may be related has been entertained by several scholars, see among other Anderson (1993) and (2003) and Fontana (1996).

“They told me: ‘we shall kill him and shall bring his head to you.’”, *StBoT* 24 ii 24-26
(=Otten, 1981).

In all three sentences we find clitic pronouns in P2: *-mu* 1sg.obl, *-an* 3sg.acc, and *-ta* 2sg.obl. In addition, (21b) and (21c) also feature the direct speech particle *-wa(r)*, and (21b) contains one of a set of so-called local particles, *-kan*. Other frequently found items in P2 include the modal particle, the reflexive and some sentence connectives.

4.1.1. Slots in the clitic chain

Below I give a chart of Hittite P2 clitics and their position relative to each other (see Luraghi 1990a and 2000 for further details):

- a. sentence connectives and conjunctions: *-(y)a-*, coordinator; *-ma-*, *-a-*, adversative particles, *man-* modal particle (which also has a variant that can occur sentence-initially and host other enclitics; it can co-occur with the connective *-ma-*);
- b. *-wa(r)-*, direct speech particle;
- c. nominative or accusative of third person pronoun singular or plural;
- d. oblique forms of first and second person singular and plural or dative of third person singular or plural (in the plural dative enclitic pronouns normally precede possible nominative or accusative enclitics);
- e. *-z(a)-*, reflexive particle;
- f. *-kan*, *-(a)sta*, *-san*, *-an*, *-(a)pa*, so-called local particles.

Regarding connectives in slot (a), it must be noted that they are in complementary distribution with the sentence introducer *nu* (and with two other such connectives, *ta* and *su*, mainly found in Old Hittite). When present, the latter particle hosts clitics, as in (21a) and (21c). I return to the

prosodic status of this particle below, sec. 4.1.2. The modal particle *man* figures in slot (a) mainly on account of its non-co-occurrence with *nu*, but, as noted, it can co-occur with the enclitic connective *-ma*, and it can stand clause initially and host other clitics. Thus, P2 cliticization in the case of *man* seems to be a recent phenomenon as the particle retains a double status: in practice, we can regard it as a simple clitic, rather than as special one as other P2 clitics. When standing in initial position, the particle may well have been proclitic, similar to the connective *nu*, as discussed in sec. 4.1.2.

The direct speech particle *-wa(r)-* in slot (b) derives from a grammaticalized form of a verb of speech, cognate to the root of English *word* and Latin *verbum* (Luraghi 1996). It does not have clitic counterparts in the other ancient Indo-European languages; however, the tendency for verbs of speech to occur after the first word inside a stretch of reported direct speech is well attested.

Compare Latin:

- (22) *Hic ... “Nihil,” inquit, “de eorum sententia dicturus sum ...*
 DEM.NOM.M nothing:ACC tell:PRS.3SG of 3PL.GEN opinion:ABL say:PTCP.FUT.NOM be:PRS1SG
 “This one said: ‘I am not going to talk about the opinion of those ...’” *Caes. G. ...*

Comparative evidence points toward late grammaticalization and cliticization of this particle, which must have changed into a P2 clitic after P2 had already been established as the locus of enclisis at least for pronouns, which occur in P2 in the other ancient IE languages as well.

The grammaticalization of this verb form into a particle casts interesting light on clitic climbing as attested with serialized motion verbs. An example is *paiueni* ‘we go’ in (21b). In example (21), I have not given an independent translation of this verb form. Indeed, this verb functions as a sort of auxiliary with respect to the second verb form in the same sentence, *kuennumeni* ‘we kill’.

Remarkably, it hosts the third person accusative clitic *-an*, which syntactically belongs with the second verb and is its direct object, giving rise to a construction which has been shown to reflect a

process of serialization of motion verbs (Luraghi 1993). Serialized motion verbs can occur in first position as here or in ‘post-P2’ or ‘post-Wackernagel’ position, and be preceded by the sentence introducer *nu* and the clitic chain, as in:

(23)

(a) *nu= mu= kan AMA- ŠU menahhanda para naista*

CONN 1SG.OBL PTC mother-his against towards send:PRET.3SG

(b) *n= as= mu uit GÌR.MEŠ-as katta haliyattat*

CONN 3SG.NOM 1SG.OBL come PRET.3SG foot-PL:D/L.PL down fall: PRET.3SG

“He sent his mother to me: she fell to my feet.” *AM* 70.28-29 (Goetze: 1933);

In (23), *uit* ‘s/he came’ follows the clitic chain that hosts the complements of the verb *haliyattat*, and is placed in the first available position after the clitic chain. Similar to the modal particle *man*, serialized motion verbs have sentence initial and (post-)P2 variants and thus qualify as simple clitics; a possible further step in their grammaticalization process (if the Hittite language had not disappeared) might have been to develop invariable forms and become special clitics, as the particle *-wa(r)-* (for further details and more examples see Luraghi 2000: 49-51; on the possible meaning of serialized motion verbs, which is still debated, see Hoffner, Melchert 2008: 324-329).

Slots (c)-(e) of the clitic chain host pronouns and reflexives; some of them have precise cognates among P2 enclitic pronouns of other IE languages. Slot (f) hosts the so-called local particles, a category which has been puzzling Hittitologists for a long time as they have no straightforward counterpart in the other IE languages. In New Hittite, local particles (notably the particle *-kan* as the other are rather typical of the older language) tend to express aspectual meanings; in Old Hittite, however, their original function as spatial deictic was still clear. As discussed especially by Boley (see Boley 1989 and 2000), these particles resemble Romance adverbial clitics: noting that *-asta* tends to co-occur with source expressions, Boley compares it to Italian ablative clitic *ne* ‘from

there' (cf. example 9); similarly, *-san*, which often co-occurs with locatives, can be compared with Italian locative clitic *ci* (cf. example 10). In Luraghi (2001) it is argued that these deictics are etymologically related to the PIE preverbs, which, similar to what happened in Sanskrit, were renewed by phonologically heavier adverbs. Note that the change from spatial deictics to perfectivity markers is consistent with this scenario and indeed represents a common evolution of preverbs in the Indo-European languages and cross-linguistically (see Pinault 1995, Lazard 1995).

4.1.2. Trends in clitic clustering

Summarizing, we find two types of item in the Hittite initial chain: (a) sentence connectives or other particles that indicate sentence type (modal particle, direct speech particle), (b) personal pronouns and deictics. Interestingly, the items in (a) precede those in (b): far from being a feature of Hittite, this is the same distribution found in the other ancient IE languages, such as Sanskrit and Ancient Greek (see sec. 4.2 for ancient Greek; on Sanskrit, see Hock 1996 and Schäufele 1996).¹⁷ The two types of item have different scope; accordingly their placement in P2 has different motivations: items in (a) have the whole sentence in their scope, so their occurrence at the sentence boundary is not surprising. Indeed, they occur as early in the sentence as possible for enclitic elements, and attach to the left hedge of their host. In terms of Klavans (1985), sentence particles attach to the same host both syntactically and phonologically. Pronouns and deictics have the VP in their scope. They are functionally similar to Romance clitics, but tend to move to the left because, as pointed out in the case of clitic climbing in Italian, they refer to non-focal, shared information. Again, using the structural analysis in Klavans (1985), P2 pronominal and spatial deictic enclitics have the VP as

¹⁷ Hock (1996) pursues a prosodic analysis of Wackernagel's law in Vedic, and proposes a templatic approach to the order of clitics in clusters, whereby accented and unaccented forms alternate. As Schäufele (1996) points out, it may well be the case that Hock's templatic approach is fundamentally right; however, the fact remains that sentence particles precede pronouns.

their structural host, but take the sentence as their phonological host.

The fact that pronominal clitics in Hittite do not attach to the verb and hence do not tend to become verbal affixes does not indicate a lower degree of grammaticalization. As argued in Luraghi (1998), clitic chains, typical of Hittite and in general of Anatolian at all language stages, tend to acquire more members, become longer and more frequent after Old Hittite. This happens through a variety of developments, such as increasing obligatoriness of the reflexive particle, which takes over many of the functions of a reflexive middle, grammaticalization of the spatial particle *-kan* as an indicator of telicity, demise of NP based clitic possessives, which are substituted by P2 dative clitics (an example of the latter usage is *-mu* in 23b, to be compared with the NP hosted possessive *-ssi* in example 3).

When a constituent needed to be left dislocated, clitic chains served the purpose of isolating it by creating a prosodic boundary between the left dislocated constituent and the rest of the sentence. In case no constituent needed to be left dislocated, the sentence introducer *nu* was needed just for the purpose of hosting clitics. This particle was most likely itself unaccented, as shown by its tendency to become phonologically fused with following enclitics when they started with a vowel: *nu+as = nas*. Note however that, when occurring without clitics, it was separated in writing from the next word: in other words, possible proclisis was not noted in writing, while enclisis always was.¹⁸ In sentences introduced by *nu* and a clitic chain, one cannot speak of a properly accented constituent in first position: the whole clitic group is rather likely to be a weak element, perhaps even proclitic to the next accented word.

In conclusion, I would like to suggest that Hittite featured two types of clitic group: one, with an accented word as the host of the clitic cluster, which served some special discourse function by

¹⁸ This is also true of other writing systems, see sec. 5.

introducing discontinuity into the text,¹⁹ and another, with a weakly accented clitic group hosted by a connective (or possibly by the modal particle *man*, see above sec. 3.1.1), which appeared in context of discourse continuity. Discontinuity was introduced into the text by fronting a specific word and separating it from the rest of the clause by a prosodic boundary: this happened with left dislocated topics, but also with preposed local adverbs, initial verbs,²⁰ or other items that did not stand in their habitual position. Very frequently, such clitic clusters included the adversative connective *-ma*, whose function was precisely to indicate textual discontinuity or unexpected events (see Luraghi 1990a for a thorough discussion of word order types and sentence connectives in Hittite).

4.2. *Ancient Greek clitics*

Ancient Greek is peculiar in having an extremely free word order, mostly sensitive to discourse organization. For this reason, tendencies in clitic placement, as well as in the placement of other pre- and postpositives (see below for explanation) have been regarded as strikingly in contrast with free positioning of accented constituents. What has attracted the attention of scholars for over a century is not only the regularity for certain items to always be excluded from first position and their frequent appearance in second position, but also the fact that, in spite of this regularity, P2 clitics often seemed to ‘skip’ some initial element. Prosodic descriptions of this phenomenon reach back at least to Fraenkel’s seminal study on the relation between clitic placement and cola, units defined in prosodic terms (see Fraenkel 1932, 1933, and Goldstein 2010 for discussion). In this

¹⁹ Textual discontinuity could have a stylistic function, or it could indicate an unexpected event, or signal clauses that introduced backgrounded information, see Luraghi (1990a).

²⁰ Hittite was a fairly strict SOV language and initial verbs were quite exceptional at all language stages, see Luraghi (1990a).

section, I give a short and necessarily schematic description of clitic placement in Ancient Greek, and show how discourse factor favored the demise of P2.

4.2.1. P2 clitics outside P2

As already noted by Wackernagel, in Greek and Sanskrit certain words could be skipped by P2 clitics, with the consequence that the ensuing word counted as first. Such words were in the first place subordinators and sentence connectives (e.g. Sanskrit *athā*, see Luraghi 1990b, Hock 1996), as shown in the following examples:

(24) *kai líēn se páros g' out'eíromai oúte metallô*

and frequently 2SG.ACC before FOC NEG ask:AOR.1SG.MID NEG inquire:AOR.1SG

“In the past I have not been accustomed to inquire nor ask you.” Hom. *Il.* 1.553;

(25) *hòs tóte mén moi hupéskheto kai katéneusen*

REL.NOM once PTC 1SG.DAT promise:AOR.3SG.MID and bow:AOR.3SG

“Of old he promised me, and bowed his head thereto.” Hom. *Il.* 9.19.

Example (25) features two types of P2 items, i.e. the enclitic pronoun *moi* (1sg.dat) and the postpositive discourse particle *mén*. Such particles are never found sentence intially, hence the name of postpositive used to refer to them (e.g. Dover 1960); many of them, such as *mén*, always bear a graphic accent. As accents have been added in Greek text only at a late time, it is not clear what value one should attribute them. Limited evidence can be adduced from Mycenaean regarding the unaccented status of *dé* (Bartonek 2003: 128); note however that the distribution of postpositives in Classical Greek is not the same as the distribution of enclitic pronouns. In particular, discourse particles such as *mén*, *dé*, *gár* frequently occur between the definite article and

the noun it determines, while such placement is quite exceptional for enclitic pronouns.²¹ More in general, the tendency for such discourse particles to follow the first word in a sentence is much stricter than for enclitic pronouns, which can instead occur later, after pragmatically relevant prosodic units. Thus, a typical patterns found in Classical Greek prose is the following:

- (26) *ho gár toi paîs me ho Sáturos apédra*
 ART.NOM PTC PTC child.NOM 1SG.ACC ART.NOM Satyros:NOM run.away:AOR.3SG
 “The boy Satyrus ran away from me.” Pl. *Prot.* 310c3.

In (26) the particles *gár* and *toi* and the pronoun *me* are in two different types of P2: after the definite article (first word) and after the determiner phrase *ho paîs* (first constituent). Such usage has the consequence that a sentence may feature more than one second position, depending on how first position is defined. Very often, it is not only the article that hosts sentence particles: the double P2 can occur with all sorts of lexical items, as in:

- (27) *áneu gâr episēmou ou sphî nómos estî ékhein skêptron*
 without PTC image:GEN NEG 3PL.DAT custom:NOM be:PRS.3SG have:PRS.INF staff:ACC
 “Indeed, it’s not their custom to carry a staff without an image.” (Hdt. 1.195.2)

²¹ See Taylor (1990: 120). Dative enclitic pronouns can be placed between the definite article and their noun, but apparently only when they function as possessive, and can therefore be considered part of the NP, see Taylor (*ib.*) and Goldstein (2010: 93-95) for discussion. In cases in which *dé* is hosted by a proclitic, such as the definite article or the negation, Goldstein argues that it must be regarded as proclitic to the next accented word. In the light of the discussion in sec. 4.1.2 regarding possible clitic groups hosted by proclitics in Hittite, one could equally argue for similar proclitic hosts in Greek.

As in the case of Hittite clitics discussed in 4.1, Greek clitics and postpositives arguably attach to P2 for different reasons: discourse particles are sentence introducers, and they occur at the leftmost hedge of their host; on the other hand, enclitic pronouns move there because of their low discourse saliency, but, contrary to Hittite, they can also occur elsewhere (Luraghi 1990b). Now, the interesting issue is for what reason they should do this.

As argued in Luraghi (1990b) and discussed at length in Goldstein (2010), parts of sentences that occur to the left of the domain of cliticization as in (27) are left dislocated for specific pragmatic reasons: in this case, Goldstein (2010: 154) shows that the phrase *áneu gàr episēmou* is focused. Goldstein further argues convincingly that what can be separated cannot always be defined in syntactic terms as a constituent, but is best regarded as an intonational phrase.

4.2.2. Prosodic boundaries inside sentences

Similar to Hittite, in Ancient Greek, too, P2 enclitics and postpositives created a boundary inside a sentence by isolating a part of it through a cluster of prosodically deficient elements. In Luraghi (1990b) it is further argued that the occurrence of postpositives and enclitics in more than one position within the same sentence had the effect of creating more than one boundary. In this section, I show how such sentence internal boundaries operate by examining a frequently cited passage which has never been explained in a satisfactory manner.

Consider the position of the clitic *min* (3sg.acc) in (28):

(28)

- (a) *eselthoûsan* *dè kai tithêisan* *tà* *heímata*
 go.out:PRT.AOR.ACC.F PTC and put:PTCP.PRS.ACC.F ART.ACC.PL clothe:ACC.PL
ethêito *ho* *Gúgēs*
 see:IMPF.3SG.MID ART.NOM Gyges:NOM

“When she had come in and was laying aside her garments, Gyges saw her; (when she turned her back upon him to go to bed, he slipped from the room.)”

(b) *kaì hē gunē eporâi min exiônta*

and ART.NOM woman.NOM see:PRS.3SG 3SG.ACC go.out: PTCP.PRS.ACC

“The woman glimpsed **him** as he went out. (But when she realized what her husband had done, though shamed, she did not cry out or let it be seen that she had perceived anything, for she meant to punish Candaules)” Hdt 1.10.2.

Two different explanations have been given for the occurrence of *min* in (28) viz. that the verb *eporâi* is given prominence, or that the NP *hē gunē* is extraposed for the sake of creating suspense. Goldstein (2010: 15, 145) discusses both. Regarding the first interpretation, he points out that, by focusing on the possible pragmatic status of the verb, it fails to recognize that the important fact about this occurrence is that the intonational phrase *hē gunē* is extraposed; the clitic then is placed at the left hedge of the second intonational phrase and it is just by coincidence that the actual word that hosts it is the verb.

Proponents of the second explanation claim that the fact that the woman saw Gyges was unexpected, and that by left dislocating the constituent and delaying the P2 clitic Herodotus creates suspense in the narration. Goldstein counters this explanation by quoting a similar passage:

(29) *kaì hē Ámēstris punthánetaí min ékhousan*

and ART.NOM Amestris:NOM learn:PRS.3SG 3SG.ACC have:PRT.PRS.ACC

“(But as he [*sc.* the king] could not move her, he gave her the mantle; and she, rejoicing greatly in the gift, went flaunting her finery.) Amestris heard that she had **it**. (But when she learned the truth, it was not the girl with whom she was angry. She supposed rather that the girl's mother was guilty and that this was her doing, and so it was Masistes' wife whom she plotted to destroy).” Hdt. 9.110.2.

In (29), Goldstein argues, the suspense effect is not needed as the fact that her husband's lover had the mantel was exactly what she expected. However, this passage must be considered in a wider context, as I will argue after discussing (28).

In the first place, the occurrence of *min* in the above examples also deserves to be considered in a wider framework, and to be explained, as opposed to the possible occurrence of a null object. In very much the same way as we have seen above for Latin (sec. 3.3), referential null objects occur in Ancient Greek. Indeed, as discussed in Luraghi (2003), the above examples represent a typical context in which discourse conditioned null objects could occur: this is made clear in (28a) and (28b). In this passage both participants referred to, the woman (Candaules' wife) and Gyges are topical; however, in the first part of the example we find a null object referring to *hē gunē* while in the second part *ho Gúgēs* is resumed by the clitic *min*. Note that in both sentences predicative constituents referring to the direct object also occur. As indicated in Luraghi (2003), the difference is indeed connected with the second event being unexpected, not simply because the woman did not expect to see someone looking at her while she was undressing, but rather because her husband Candaules, who had arranged for Gyges to see her, had not thought of the consequences. In this sense, the event in (29) is also unexpected, because the king's lover who insisted on receiving the mantel had miscalculated the possible consequences. Note that the following context is remarkably similar in the both passages; in both cases, the reaction of the woman who unexpectedly discovered someone else's misdeed had quite dramatic consequences on the person who suffered them.

Thus, not only the position, but the very occurrence of the clitic *min* depends on the unexpectedness of the event. In both examples the subject, which refers to the person who discovers something that should have been concealed, is left dislocated, the occurrence of the clitic after the verb identifies the two parts of the sentence as distinct intonational phrases, while in normal conditions there would have been no need to split up the sentence into two intonational phrases. The clitic group constituted by the verb and the clitic stands out as indicating a boundary inside the

sentence, which signals that the event marks a turning point in the life of a previously mentioned highly topical participant.

4.2.3. The demise of Wackernagel's Law

As well known, Modern Greek features pronominal clitics hosted by the verb. According to some scholars, including Wackernagel (see Goldstein 2010: 11-13 for further reference), the first traces of the tendency for clitics, mostly pronominal, to leave P2 and gravitate around the verb can be detected in Classical Greek prose already. According to Goldstein (2010), this is not the case. In his interpretation of the data, clitics can be shown to attach to the leftmost word in an intonational phrase; when they are not sentence initial, it is because one or more other intonational phrase(s) have been 'skipped', or extraposed to the left for pragmatic reasons, being topical or focal.

It may well be the case that Goldstein's explanation is correct: however, as I have shown above, the consequence was that possible locations for P2 clitics occurred at several points in a sentence. This tendency is specular to what we have seen in sec. 4.1 regarding Hittite clitic chains: rather than being strictly confined to a uniquely defined P2, Greek clitics could be scattered in a sentence, at a potentially unlimited number of intonational breaks. Note further that the status of pronominal clitics as 'special' clitics is far from being clear: I cannot pursue this issue here, suffice it to mention that the third person clitic *min*, shown in examples (28) and (29) above, does not occur in Attic prose, which constitutes the bulk of classical literature. There, third person is indicated by non-nominative weak forms of the former demonstrative *autón*, which was most likely a simple clitic, similar to Latin *eum* (see sec. 3.3). In addition, as discussed by Goldstein, some pronominal forms must be taken as proclitic to explain certain metrical facts, thus apparently indicating a double status as to phonological liaison.

In conclusion, Classical Greek offers a picture in which particles and weak forms enjoyed a remarkable freedom in their position, once prosodic conditions were satisfied, thus serving a variety of discourse functions, favored by the high degree of word order freedom, including possible

occurrence of discontinuous constituents of all sorts. Similar to Latin, Greek also displays increasing grammaticalization of word order, ultimately motivated by increasing configurationality (see Devine, Stephens 1999, Luraghi 2010a), with accompanying loss of null objects, and the development of a highly grammaticalized system of verb based clitics. The latter exhibit features of head marking strategies in Modern Greek, as indicated by frequent clitic doubling (see Joseph 2002 with the references therein and Janse 2000 for a comparison between the development from Latin to Romance and from Ancient to Modern Greek). The picture offered by Classical Greek does not yet indicate such late developments: the freedom of clitic placement was a consequence of the high extent to which Greek word order was discourse conditioned. Remarkably, as pragmatic factors gave way to a more syntactically conditioned sentence structure, clitics, which were not all clustering in the same P2 position, had so to speak nowhere to go, and ended up being attracted to the verb and form a structurally defined VP with it.

5. Differences between enclisis and proclisis?

Observing orthographic representation of proclitics and enclitics in the Romance languages, Benincà and Cinque (1983) suggest that the fact that enclitics are often attached to their host graphically whereas proclitics are not may reflect some difference in the relation between the host and the clitic based on the direction of liaison. Remarkably, such graphic differences are not only found in the Romance languages, but also in various other completely unrelated graphic systems, such as that of Hittite and, possibly, Mycenaean Greek:²² it is frequently the case that in graphic representation enclitics are spelled as one word with their host, while proclitics are not. In the light of the

²² According to Adams (1996), even some Late Latin texts provide evidence for enclisis of pronouns by not separating them in writing from the preceding word (mostly the verb).

discussion of possible proclitics hosting enclitics in sec. 4.1, the possibility that the two types of clitics indeed present prosodic differences deserves to be addressed.

Among the evidence for closer binding between a host and an enclitic in Romance, Benincà and Cinque mention the impossibility of omitting a direct object clitic in the second conjunct of coordinated clauses. Recall that this type of omission was obligatory in Latin, and still occurred, even though not systematically, in Old Italian. In Modern Italian, omission can sometimes occur. Acceptability judgements vary among speakers, but in any case, omission is restricted to V coordination in cases in which the object in the first clause is a clitic. Compare the following example:

- (30) *L'esistenza contemporanea non è minacciata ... dall'impeto delle acque profonde, ma dalla bonifica che le inaridisce e prosciuga,*
the existence contemporary not is threatened by+the thrust of+the waters deep but by+the
drainage which them(cl) wither and dry.up
“Contemporary life is not threatened by the outburst of waves from deep waters, but by the drainage which dries them up.” C. Magris, *Dietro le parole*, Garzanti, Milano, 1978.

Such type of omission would be impossible with enclitics:

- (31) **la bonifica che minaccia di inaridirle e prosciugare / *di inaridire e prosciugarle*
the drainage which threatens to wither them (cl) and dry.up / to wither and dry.up them (cl)

In this respect, enclitics behave as bound morphemes: each verb requires an overt enclitic, while proclitics can have scope over two coordinated verbs.

Benincà and Cinque further mention sporadic occurrences of adverbs intervening between a proclitic and the verb from various Romance varieties including Old Italian (note that the better

known occurrences of separation of a clitic from its verb in Old Spanish are also limited to proclitics, see Fontana 1996 and Wanner 1996 for the examples):

(32) *e se ella il mi pur crede*

and if she it(cl) me(cl) even believes

“If she believes what I’m saying.” (Boccaccio *Decameron* viii.6).

In addition, proclitics can be coordinated in French and Romanian, while enclitics cannot:

(33) *Je lui et vous ferais un plaisir.*

I him(cl) and you(cl) will.do a favor

“I will do him and you a favor.”

The evidence reviewed thus far points toward some higher degree of autonomy of proclitics. However, only adducing evidence from pronouns has the obvious risk that one might confuse properties of anaphoric reference with properties of clitics. For this reason, the evidence provided by Hittite and other languages in which otherwise proclitic items can host enclitics is very important. In such languages, the clitic cluster can be hosted by a proclitic, but it looks doubtful that it can constitute an autonomous intonational unit, in spite of having some autonomy: most likely it leans to the left hedge of the intonational unit on its right. Thus, the initial proclitic rather than receive an accent from the enclitics, might better be described as transmitting its liaison properties to the enclitics, so that the whole cluster becomes proclitic.

It could then be the case that initial chains of enclitics, when hosted by an unaccented clause introducer, become proclitic. As a consequence, they can no longer be said to be placed in P2, but

should rather be viewed as standing in P1.²³ P1 placement is an extreme consequence of the tendency, repeatedly remarked for clitics, to occur early in a sentence. In languages in which clitics tend to cluster in P2, recurrent clitic chains can be exploited for various discourse functions, such as that of isolating extraposed words or constituents by means of a prosodic boundary constituted by a string of unaccented items. Association of clitic clusters with emphasized, focused, topicalized constituent and the like emerges in usage as a recurrent pattern; its entrenchment in the speakers' cognition creates the need for a different sentence pattern, in which no constituent is isolated in such a way. The introduction of a proclitic clause introducer has the effect of extracting the internal boundary formed by unaccented items from the clause, and create a clause type in which, in spite of the occurrence of clitics, no constituent is isolated.

6. Conclusions

In this chapter I have given a usage based account of various characteristics of clitics, including their very existence and tendencies in their placement rules. I have shown that discourse organization plays an important role in determining peculiar properties of clitics. Regarding clitic placement, it is still worth remarking that the two main strategies described, i.e. for clitics either to cluster around a specific constituent or to occur early in a sentence, typically in P2, can be explained as owing to competing motivations. In the first case, clitics tend to lean phonologically to the constituent that they have in their scope, or to which they belong syntactically, as in the case of focalizers (sec. ...) and of pronominal clitics, such as those of the Romance languages. In the second case, clitics tend to occur early in a sentence because of their function in discourse: for example,

²³ It is worth noting that Agbayani and Golston (2010) reach a similar conclusion within a formal framework, referring to syntactic cliticization of P2 enclitics in Ancient Indo-European languages.

pronominal clitics carry old information, already introduced into the discourse, and serve referent tracking. So clitic placement may have different motivations, which drive clitics to attach to different hosts; competing motivations can also be the cause of change in clitic placement.