Abstract

In this paper we discuss second position clitics in ancient Greek, which show a remarkable ability to break up syntactic constituents. We argue against attempts to capture such data in terms of a mismatch between c-structure yield and surface string and instead propose to enrich c-structure by using a multiple context free grammar with explicit yield functions rather than an ordinary CFG.

1 Introduction

Second-position (2P) clitics have proven notoriously challenging for syntactic theory, because their distribution requires reference to both syntactic and prosodic constituents, as illustrated by the following example from ancient Greek ('=' marks prosodic dependence):

(1) *2P distribution*

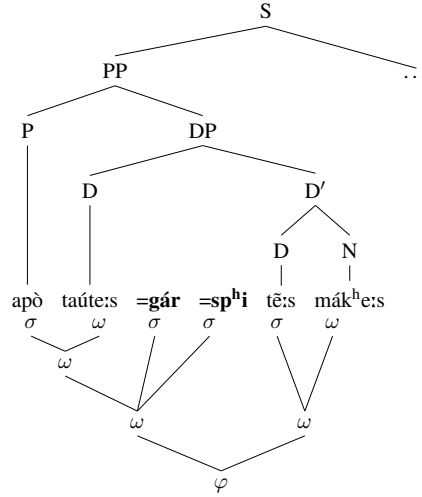*(apò taúteːs)$_\omega$=**gár**=**sp^{h}i** tẽːs mák^{h}eːs* ...
from MED.F.GEN.SG=EXPL=3PL.DAT ART.F.GEN.SG battle.F.GEN.SG
kateúk^{h}etai ho kéːruks
pray.PRES.IND.MP.3SG ART.M.NOM.SG herald.M.NOM.SG
ho At^{h}eːnaĩos háma te
ART.M.NOM.SG Athenian.M.NOM.SG together.ADV CONJ
At^{h}eːnáioisi légoːn
Athenian.M.DAT.PL speak.PTCP.PRES.ACT.M.NOM.SG
gínest^{h}ai tà agat^{h}à kaì
happen.INF.PRES.MP ART.N.ACC.PL good.N.ACC.PL CONJ
Plataieũsi.
Plataean.M.DAT.PL
'Since this battle ..., the Athenian herald prays that good things befall the Athenians and Plataeans together, when the Athenians conduct their sacrifices at the festivals that occur every four years.' (Hdt. 6.111.2)[1]

The enclitics *gár* and *sp^{h}i* are hosted by the first prosodic word in the clause, (*apò taúteːs)$_\omega$* 'from this'. As the preposition and demonstrative do not form a syntactic constituent, a mismatch between syntax and prosody results:

---

[1]The data for this paper come primarily from Herodotus' *Histories*, a corpus of 189,489 tokens written in the Ionic dialect of classical Greek in the 5th century BCE. Section 2.2 introduces some metrical data from fifth century BCE Attic drama. Our transliteration of the Greek is graphic, not phonological/phonetic.

(2) *Syntax-prosody mismatch*

```
                          S
              _____/ _____
             PP                       …
         ___/  \___
        P          DP
        |      ___/  \___
        |     D          D′
        |     |        _/  \_
        |     |       D      N
        |     |       |      |
      apò  taúteːs =gár =spʰi tẽːs mákʰeːs
       σ     ω      σ    σ    σ     ω
        \___/
          ω
           _____   ___/
                  \ /
                   ω              ω
                    \____   _____/
                         \ /
                          φ
```

This mismatch raises fundamental questions about the architecture of the grammar. First, what is the division of labor between syntax and phonology—can syntax see phonological properties? Below we argue that no information passing between syntax and prosody is needed beyond the ordinary interaction of projections in the LFG architecture, which means that a well-formed sentence must simultaneously satisfy syntactic and prosodic constraints. Syntax must therefore position the clitic where it gets a prosodically acceptable host.

Second, what are the capacities of c-structure and f-structure? The root of the problem in example (2) is that clitics appear in surface positions where they cannot be assigned a grammatical function. This inability is a direct consequence of the linguistic and formal differences between c- and f-structure. Linguistically, c-structure deals with word order and constituency, whereas f-structure deals with abstract syntactic relations. Formally, c-structure can only handle phenomena within the locality domain of a context-free grammar (CFG), i.e. the one level tree corresponding to a rule whereas f-structure can handle phenomena at an unbounded distance. What this amounts to is the claim that there are no non-local word order or constituency facts. But 2P clitics involve precisely non-local constituency.

To deal with non-local constituents, previous accounts have surrendered core assumptions of LFG by introducing idiosyncratic constituents, such as CL and CCL (i.e. syntactic categories 'clitic' and 'clitic cluster', Bögel et al. 2010, Ćavar & Seiss 2011, Lowe 2011); by relying on cross-derivational comparison (i.e., Output-Output Correspondence, Lowe 2015) under Optimality Theory; by permitting mismatch between the c-structure and the prosodic or syntacic string (Bögel 2015, Lowe 2015); or by allowing prosodic markers into syntax Bögel et al. (2009, 2010). We return to some of these proposals in greater detail below in section 4.

By contrast, we argue that the best way to capture the empirical facts and to maintain the spirit of the LFG architecture is to modify the division of labor between the c- and f-structures. Specifically, we increase the power of the c-structure
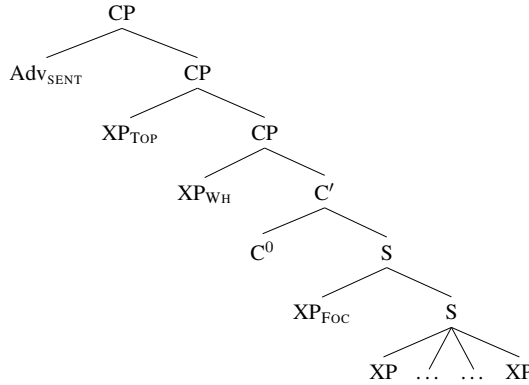
Figure 1: AG clause structure

by moving to a 2-multiple context-free grammar (2-MCFG). This move enables us to capture two crucial insights into the behavior of 2P clitics in ancient Greek. First, there are no clitic specific syntactic rules. Second, 2P clitics do not occupy dedicated c-structure positions. This results from the prosaic fact that syntactic constituents need not map onto identical prosodic constituents. 2P clitics only subcategorize for a syntactic domain and host properties.

2 Data

2.1 Background on ancient Greek

Ancient Greek is one of the earliest attested Indo-European languages. It is a language with rich nominal and verbal morphology and "free" word order, which includes remarkable discontinuities (Devine & Stephens 1999, Agbayani & Golston 2010a). Whether or not Greek has an underlying configurational word order has been debated for over a century, with the main contenders being OV and VO. What is clear is that Greek relies heavily on surface word order for encoding pragmatic properties of the clause, such as information structure (Dik 1995). Since no correlation between surface position and grammatical function has been demonstrated, we assume an exocentric S constituent for the basic clause. Various phrases encoding information structure functions can be adjoined to CP or to S, as shown in Figure 1, which is based on and deviates slightly from the clause structure proposed by Goldstein (2016).

Beginning at the top of the tree, adverbials with sentential scope adjoin at the highest level, with topicalized phrases occuring next. Both phrases are adjoined to CP. We assume that *wh*-words occupy Spec,CP. A focus projection can be adjoined to S, which broadly speaking encodes contrastive focus (Goldstein 2016 refers to it as NON-MONOTONIC FOCUS). This gives us the c-structure rules in (3).[2]

---

[2] The order of adjunctions to CP in Figure 1 does not follow from these rules: a sentential adverb

(3) *C-structure rules*

<div align="center">

Clause-level rules

</div>

$$\begin{array}{lclcl}
\text{CP} & \rightarrow & \text{XP} & & \text{C}' \\
& & (\uparrow\ \text{UDF})=\downarrow & & \uparrow=\downarrow \\
\text{C}' & \rightarrow & \text{C}^0 & & \text{S} \\
& & \uparrow=\downarrow & & \uparrow=\downarrow \\
\text{S} & \rightarrow & \text{XP*} & , & \text{V*} \\
& & (\uparrow\text{GF})=\downarrow & & \uparrow=\downarrow
\end{array}$$

<div align="center">

Adjunction to clausal categories   *Lexical phrases*

</div>

$$\begin{array}{lclllclll}
\text{CP} & \rightarrow & \text{AdvP} & \text{CP} & \text{PP} & \rightarrow & \text{P} & & \text{DP} \\
& & \uparrow=\downarrow & \uparrow=\downarrow & & & \uparrow=\downarrow & & (\uparrow\ \text{OBJ})=\downarrow \\
\text{CP} & \rightarrow & \text{XP} & \text{CP} & \text{DP} & \rightarrow & \left(\begin{array}{c}\text{D}\\ \uparrow=\downarrow\end{array}\right) & & \text{NP} \\
& & (\uparrow(\text{GF}))=\downarrow & \uparrow=\downarrow & & & & & \uparrow=\downarrow \\
& & (\uparrow_\sigma\ \text{DF}) = \text{TOPIC} & & \text{NP} & \rightarrow & \text{A*} & , & \left(\begin{array}{c}\text{N}\\ \uparrow=\downarrow\end{array}\right) \\
\text{S} & \rightarrow & \text{XP} & \text{S} & & & (\uparrow\text{ADJ}) \in \downarrow & & \\
& & (\uparrow(\text{GF}))=\downarrow & \uparrow=\downarrow & \text{AdvP} & \rightarrow & \text{Adv} & & \\
& & (\uparrow_\sigma\ \text{DF}) = \text{FOCUS} & & & & \uparrow=\downarrow & &
\end{array}$$

The most prominent feature of these rules is the absence of any reference to clitics. In contrast to other accounts (see section 4 below), our analysis posits no clitic-specific syntactic rules. So the grammar handles pronominal clitics just as it does their stressed counterparts, and clitic discourse particles with sentential scope are treated just like sentential adverbials.

Determining the category of second-position clitics is no easy matter. Several recent LFG analyses rely on a category CL (for 'clitic'), while minimalist analyses often posit a category between that of a syntactic head and phrase (e.g., D/P for pronominal clitics). Here we take a different tack and analyze 2P clitics in AG as non-projecting heads (Toivonen 2003). However, unlike Toivonen, we assume that what is special about non-projecting words is simply that they do not project, which correctly predicts that they cannot, e.g., be the targets of adjunction.

2.1.1 Clitics versus postpositives

The philological literature standardly divides the inventory of second-position items into two classes: clitics and postpositives (Chandler 1881, Fraenkel [1933] 1964, Probert 2003). Second-position items without an orthographic accent are clitics (e.g., the third person singular accusative pronoun *min*), while those with an orthographic accent are postpositives, e.g., the modal particle *án*. The idea behind this division seems to be that of true phonological clitics versus "syntactic clitics," that is, words that despite bearing an accent are nevertheless restricted to second position (Devine & Stephens 1994, 303, 352, Dik 1995, 37–38, Lowe 2014).

Although we believe that the graphic distinction found among postpositives and enclitics does reflect something prosodically real, we reject the traditional view

---

must always go higher than a topic. We assume this is related to scope, but it is notoriously difficult to capture adjunction scope in LFG.

for two reasons. First, metrical evidence demonstrates that both postpositives and enclitics exhibit phonological dependence on a host (see further Goldstein 2016, 52–53, 61–65). So they are not syntactic clitics. Second, there is no distributional difference that correlates with the presence or absence of a graphic accent (cf. Taylor 1990, 119, Fortson 2010, 161).

Instead, we argue that the prosodic distinction between postpositives and enclitics is limited to secondary stress assignment. When enclitics and postpositives incorporate with a prosodic word (i.e., their host), they uniformly project a secondary prosodic word (see Anderson 2005 for an overview, including earlier literature):

(4) *Recursive prosodic word*

$$\omega$$
$$\omega \qquad \sigma$$

Host Enclitic/Postpositive
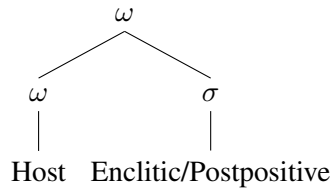
The secondary prosodic word can trigger a secondary stress, the position of which is determined in one of two ways. The first is via a secondary accentual calculus, whose precise details do not concern us here. Suffice it to say that enclitic incorporation can trigger secondary stress on the host, the clitic, or none at all:

(5) *Secondary stress patterns*

    i. *Secondary stress on the host*

        $(ánt^h ro{:}poi)_\omega+tines \qquad\qquad \rightarrow ((ánt^h ro{:}poí)_\omega=tines)_\omega$
        man.M.NOM.PL+INDF.C.NOM.PL

    ii. *Secondary stress on the clitic*

        $p^h íloi+tines \qquad\qquad \rightarrow ((p^h íloi)_\omega=tinés)_\omega$
        friend.M.NOM.PL+INDF.C.NOM.PL

    iii. *No secondary stress*

        $(pánta)_\omega+sp^h i \qquad\qquad \rightarrow ((pánta)_\omega=sp^h i)_\omega$
        all.N.NOM/ACC.PL+3PL.DAT

The crucial point for our account is that the surface effect of the incorporation of an enclitic is variable. It is this variety that distinguishes enclitics from postpositives. Incorporation of the latter always triggers secondary stress, which is uniformly located on the postpositive itself:

(6) *Fixed secondary stress*
    $(taúte{:}s)_\omega+gár \rightarrow ((taúte{:}s)_\omega=gár)_\omega$

Secondary stress will always occur on the postpositive, regardless of the prosodic shape of the host.

Before moving on, we call attention to one context in which the combination of a host plus enclitic does not project a secondary prosodic word:

(7) *hoi+gár*   → *(hoi=gár)*$_\omega$
     ART.M.NOMPL+EXPL

Examples of this type are remarkable because the host of the postpositive is a sub-prosodic word. The definite article is standardly assumed to be proclitic, but together with the postpositive it projects a prosodic word. Evidence that a string such as *hoi=gár* forms a prosodic word comes from its ability to host clausal clitics (see example 8i below; cf. also Goldstein 2016, 76–78). Similar behavior is known from Bilua, a Papuan language of Solomon Islands (see further Anderson 2012).

2.1.2   Clitic domains

The clitic lexicon of AG is larger than that of any other archaic IE language. It encompasses pronouns, verbs, conjunction, and discourse and modal particles. There is no single "second" position in which they all occur (cf. Hale 1987a,b on Sanskrit clitics).[3] Instead, clitics subcategorize for particular syntactic domains, as detailed in Table 1.[4]

DOMAIN	MEMBERS
SENTENCE	{*dé, mén*}—*gár*—*õːn*—{*déː, dẽːta*}
CLAUSE	*án*—{*kote, kou, koː, koːs, kéː(i)*}—*ára*—ACC—DAT—{*eimí, pʰeːmí*}$^?$
PHRASE	*te*—{*dé, mén*}—*ge*

Table 1: Clitic domains and the internal ordering of their members

Sentence clitics are invariably discourse connectives marking intersentential relations: we assume they are $\widehat{\text{Adv}}$. Clausal clitics realize grammatical features of the clause: they can be $\widehat{\text{Adv}}$, $\widehat{\text{D}}$ and $\widehat{\text{V}}$. Phrasal clitics realize grammatical features of sub-clausal XPs, and will be ignored here. A clitic can be a member of more than one domain. The clitics *dé* and *mén*, for instance, exhibit both sentential and phrasal scope (as in example 10 below, where it scopes over a topicalized phrase).

Clitic domains mirror clitic scope: CP for sentential clitics, S for clausal clitics, and sub-clausal XPs for phrasal clitics. This follows from their syntactic categories, since sentential adverbs (including clitic $\widehat{\text{Adv}}$) must be adjoined to CP whereas argument DPs/$\widehat{\text{D}}$s must be daughters of S to get the correct GF assigned. When there is no topicalized or focalized element (adjoined to CP or S respectively), sentential clitics immediately precede clausal clitics. When there is a topicalized or focalized element, then SPLAYING results, that is, the sentential clitic and clausal clitic are not adjacent (see example 10 in the next section). We take this to mean that in principle the whole CP with the core clause S forms one prosodic domain (an IntP); but whenever there is adjunction to CP or S for topicalization or focalization

---

[3]Some accounts have failed to incorporate this insight, e.g., Agbayani & Golston (2010b).

[4]We abstract away here from the motivation for the order of clitics within each domain, and leave open the question of whether it results from the phonology, morphology, or syntax, or some combination thereof.

purposes, this creates a prosodic break so that the core S is one IntP and the material to its left is another IntP.

Sentential and clausal clitics differ in how discriminating they are of their host. Clausal clitics are almost always hosted by prosodic words, while sentential clitics are routinely hosted by both prosodic words and sub-prosodic words:

(8) *Host variability of sentential clitics*

    i. *Sub-prosodic word host*

       *(hoi=**gár**)_ω=**me**   ek   tẽːs*
       ART.M.NOM.PL=EXPL=1SG.ACC from ART.F.GEN.SG

       kóːmeːs   paĩdes   . . . estéːsanto
       village.F.GEN.SG child.C.NOM.PL   appoint.PFV.IND.MID.3PL

       basiléa
       king.M.ACC.SG
       'For the children from the village. . . appointed me their king.' (Hdt. 1.115.2)

    ii. *Prosodic word host*

       *(tà   toiaũta)_ω=**gàr**   érga   ou   pròs*
       ART.N.ACC.PL such.N.ACC.PL=EXPL deed.N.ACC.PL NEG by

       toũ   hápantos   andròs
       ART.M.GEN.SG all.C.GEN.SG man.M.GEN.SG

       nenómika   gínesthai...
       think.1SG.PERF.ACT.IND happen.PRES.ACT.INF
       'For I have thought that not each man is capable of such deeds, but ...' (Hdt. 7.153)

In example (8i), the postpositive *gár* is hosted by the definite article *hoi*, which is a sub-prosodic word. In example (8ii), by contrast, the selfsame particle is hosted by the prosodic word (*tà toiaũta*)_ω. This variation in host selection further distinguishes postpositives from enclitics.

2.2 Second-position distribution

The basic distributional generalization is that clitics occur in the leftmost position possible modulo their lexical entries. When no material is adjoined to CP or S, sentential and clausal clitics are hosted by the first prosodic word in CP or S. They will furthermore be adjacent, with sentential clitics preceding the clausal clitics. This is illustrated by example (8i), which we repeat here for convenience:

(9) *Adjacent clitics*

    *(hoi=**gár**)_ω=**me**   ek   tẽːs   kóːmeːs*
    ART.M.NOM.PL=EXPL=1SG.ACC from ART.F.GEN.SG village.F.GEN.SG

    paĩdes   . . . estéːsanto   basiléa
    child.C.NOM.PL   appoint.PFV.IND.MID.3PL king.M.ACC.SG

'For the children from the village. . . appointed me their king.' (Hdt. 1.115.2)

The sentential clitic *gár* immediately precedes the clausal clitic *me*, because no constituents are adjoined to S. Adjunction to CP or S results in splaying if both sentential and clausal clitics are present:

(10) *Splaying*

[*tὲːn=**mèn**=**gàr*** *protéreːn* *heːméreːn*],
ART.F.ACC.SG=PTCL=EXPL previous.F.ACC.SG day.F.ACC.SG
*pánta=**sp^hi*** *kakà* *ék^hein.*
everything.N.ACC.PL=3PL.DAT bad.N.ACC.PL have.INF.PRES.ACT
'[For on the previous day], everything was bad for them.' (Hdt. 1.126.4)

The DP [*tὲːn protéreːn heːméreːn*] '(on) the previous day' is adjoined to CP. The sentential clitic *gàr* is accordingly hosted by the first prosodic word within CP. (The particle *mèn* here is a phrasal clitic that together with adjunction to CP signals the topicalized status of the DP.) The pronominal clitic *sp^hi* is hosted after the first prosodic word within S.

As mentioned above, we assume that *wh*-words occupy Spec,CP. In both direct and embedded questions, clausal clitics are hosted by the first prosodic word of the *wh*-phrase:

(11) i. *Kroĩse,* *tís=**se*** *ant^hróːpoːn*
Croesus.M.VOC.SG WH.C.NOM.SG=2SG.ACC person.M.GEN.PL
anégnoːse *epì* *gẽːn* *tὲːn*
persuade.AOR.IND.ACT.3SG against land.F.ACC.SG ART.F.ACC.SG
emὲːn *strateusámenon* *polémion*
my.F.ACC.SG campaign.PTCP.AOR.MID.M.ACC.SG enemy.M.ACC.SG
antì *p^hílou* *emoì* *katastẽːnai?*
instead friend.M.GEN.SG 1SG.DAT be.set.INF.AOR.ACT
'Croesus, what person persuaded you to stand against me as an enemy instead of with me as my ally, and campaign against my land.' (Hdt. 1.87.3)

ii. *Dareĩos* *epì* *tẽːs* *heoːutoũ* *ark^hẽːs*
Darius.M.NOM.SG on ART.F.GEN.SG 3SG.M.GEN reign.F.GEN.SG
kalésas *Hellέːnoːn* *toùs*
call.PTCP.AOR.ACT.M.NOM.SG Greek.M.GEN.PL ART.M.ACC.PL
pareóntas *eíreto* *(epì*
be.around.PTCP.PRES.ACT.M.ACC.PL ask.IMPF.IND.MP.3SG for
*kósoːi)_ω=**àn*** *k^hréːmati*
how.much.WH.N.DAT.SG=MOD money.N.DAT.SG
bouloíato *toùs* *patéras*
want.PRES.OPT.MP.3PL ART.M.ACC.PL father.M.ACC.PL
apot^hnéːiskontas *katasitéest^hai.*
die.PTCP.PRES.ACT.M.ACC.PL eat.INF.PRES.MP

'During his reign Darius summoned the Greeks who were around and asked (them) at what price they would eat their fathers after they had died.' (Hdt. 3.38.3)

In example (11i), the *wh*-interrogative hosts the pronominal clitic *se* 'you', while in (11ii), the modal particle *án* is hosted by the first prosodic word of the interrogative phrase, (*epì kóso:i*)$_\omega$. This is a significant pattern, which reveals that clausal clitics cannot be analyzed as adjoined to S (or on other analyses, TP/IP) in c-structure (if that were the case, *án* would be hosted by k^h*ré:mati*), which is what many analyses assume. Below in section 3 we show how we handle this pattern.

That prosody has the upper hand in the distribution of second-position clitics comes from examples such as the following, where the prosodic constituency of the metrical line creates the left edge of an intonational phrase:

(12) *Verse edge is intonational phrase edge*

hót-an   d'   híke:tai,   te:nikaũt' egò:
when-MOD PTCL come.PRES.SBJV.MP.3SG, then   1SG.NOM
kakòs
remiss.M.NOM.SG
(mè: drõ:n)$_\omega$=àn   éie:n
NEG do.PTCP.PRES.ACT.M.NOM.SG=MOD be.PRES.OPT.ACT.1SG
pánt^h'   hós'   àn   de:loĩ
all.N.ACC.PL so.much.REL.N.ACC.PL MOD indicate.PRES.OPT.ACT.3SG
t^{h}eòs.
god.M.NOM.SG
'When he gets here, I would be remiss if I didn't do whatever god indicates.' (Soph. *OT* 76–77)

The modal particle *án* is hosted by the third prosodic word of the clause. Crucially, it is not possible to analyze the preceding two prosodic words as adjoined phrases with either topic or focus functions. (*mè: drõ:n*)$_\omega$ is a licit host prosodic host here because the left edge of the metrical line is the left edge of an intonational phrase. So the prosodic properties of the metrical line satisfy the lexical entry of the clitic.

3   Analysis

3.1   Multiple-context free grammars

Ordinary context-free grammar (CFGs) rules conflate category formation and yield computation. A rule such as DP → D NP says both that the category DP is formed of a D and a NP, and that the yield of the resulting DP is formed by concatenating the yields of D and NP. Multiple context free grammar (MCFG) is a generalization of CFG which retains ordinary CFG productions for the expression of categorial structure, but uses explicit *yield functions* to compute the yield of the mother node

from the yields of the daughters. In effect, then, a CFG can be seen as an MCFG with concatenation as the only yield function.[5]

To allow for greater expressivity, MCFG allows yields to be *tuples* of strings. For example, we may want to say that the yield of DP is a pair (2-tuple) consisting of the yields of D and NP. This pair will then be the input to further yield functions that apply to productions with DP on the right-hand side. The start symbol of the grammar is required to yield a string.

Symbolically, we will write $\langle x, y \rangle$ to refer to the y'th component in the yield of the x'th category on the right-hand side of a production. We use semicolon (;) for concatenation and square brackets to delimit components in the yield of the left-hand side. (13) gives sample yield functions.

(13)   i. $c = [\langle 1, 1; 2, 1 \rangle]$
      ii. $s_1 = [\langle 1, 1 \rangle][\langle 2, 1 \rangle]$

c says that the yield of the mother node is formed by concatenation of the first component of the first daughter and the first component of the second daughter. s_1 (mnemonic for "split after first daughter") instead yields a pair of these two components. Notice that both rules are only defined when applied to productions with two categories on the right-hand side (since they only refer to two daughters), both of whose yields is a string (since they only refer to one component in the yield of each daughter). Since the yield of s_1 has two components, we say that it is two-dimensional.

If a daughter node is discontinuous (has dimensionality > 1), that discontinuity may be propagated to the yield of the mother node. (14) gives an example. Here the production references a particular yield function which is independently specified.

(14)   PP $\rightarrow p_2$(P DP), $p_2 = [\langle 1, 1 \rangle; \langle 2, 1 \rangle][\langle 2, 2 \rangle]$

p_2 (mnemonic for "propagate a discontinuity in the second daughter") forms a two-dimensional yield for the PP. The first component is the yield of P concatenated with the first component in the yield of DP and the second component is the second component in the yield of DP. p_2 is only defined for productions with two daughters, the second of which has a two-dimensional yield.

The split and propagation rules create two-dimensional yields but say nothing about the positioning of the two components, or what elements can intervene between them. Their function is like that of a *head* in Head Grammar (Pollard, 1984), i.e. they create a distinguished position in the yield, after which it can be split.

The idea behind our analysis is that yield functions can exploit this gap by hosting clitics in that position. This is achieved with the yield function in (15).

(15) $h_2 = [\langle 2, 1 \rangle][\langle 1, 1 \rangle; \langle 2, 2 \rangle; \langle 3, 1 \rangle]$

---

[5]Our treatment of MCFG here is necessarily brief, see Clark (2014) for an accessible introduction for linguists.

This corresponds to the wrap operation of Head Grammar: the (discontinuous) second daughter is wrapped around the (continuous) first daughter. The result is a two component yield, where the first component of the second daughter is the first component of the mother, and the second component of the mother is the concatenation of the first daughter, the second part of the second daughter and the third daughter. This yield function, then, "hosts" the first daughter within the second one, but otherwise leaves the gap in the second daughter open so that more elements can be hosted in the gap. We also need a variant of this rule which hosts the first daughter while resolving the gap in the second daughter. This is given in (16):

(16) $r_2 = [\langle 2, 1 \rangle; \langle 1, 1 \rangle; \langle 2, 2 \rangle; \langle 3, 1 \rangle]$

So far, all the yield functions we have seen deal with only one discontinuous argument. More complex situations can also arise, where there are two discontinuous arguments. In such cases we can intertwine the two to create a continuous yield for the mother node, as in (17), which takes two discontinuous arguments and one continuous one (the third).

(17) $i_1 = [\langle 1, 1 \rangle; \langle 2, 1 \rangle; \langle 1, 2 \rangle; \langle 2, 2 \rangle; \langle 3, 1 \rangle]$

As we will see, this intertwining pattern is found in examples like (11ii).

This exhausts the yield functions that we need. Importantly, c, s, h, r and i are *families* of yield functions. As we have seen, yield functions must specify how many daughter nodes they apply to and how many components they have. We leave the specification of the number of daughter nodes implicit as it is retrievable for the categorial structure. In addition, s, h and r require exactly one of the daughter nodes to have a two-component yield, and the subscript on the function designates this node. i requires two discontinuous daughter nodes; by convention, for any given i_n, this will be the n'th and $n + 1$'th daughters, so that a single parameter is enough. The number of yield functions in the grammar will therefore remain reasonable even if the branching factor is high. As we will see in section 3.2, their actual application to productions will be controlled by a prosodic *HOST* feature.

The fact that yield functions must explicitly state the number of arguments and discontinuities means that discontinuities created this way do not interact with the recursive mechanism in the categorial structure, so that the maximum number of discontinuities that a grammar permits can be read directly off the most complex yield function in the grammar. In our case, no rule outputs a yield of more than 2 dimensions, so our grammar is a 2-MCFG.[6]

It is instructive to compare this with how LFG can otherwise model syntactic discontinuities through reentrancies, i.e. multiple c-structure nodes corresponding

---

[6]This assumes that the grammar can be binarized without increasing the dimensionality, an assumption that may in fact not hold in the presence of *ill-nested* yield functions such as i, see, e.g., Kuhlmann (2013). Nevertheless, it is clear that the complexity will remain low especially given the fact that the set of potential intrudes—the admissible clitic sequences—will be finite, see section 4.4.

to the same f-structure. This does interact with the recursive mechanism of the c-structure. The result is that a given LFG may not provide an upper bound on the number of c-structure nodes corresponding to a single f-structure and hence features may in principle be transmitted across unbounded distances in the tree.

This also gives some indications of how the MCFG approach differs from linearization-based HPSG (Reape 1994; Kathol 1995), which is another attempt at dissociating categorial structure and yield computation. In linearization HPSG, the schemata that build the categorial structure do not at the same time build yields of terminals; instead, they build word order domains, over which linear precedence constraints can then be stated. A daughter's word order domain can either be *compacted*, i.e. it enters its mother's word order domain as a contiguous string yield; or it can be *domain unioned* with its mother, i.e. the elements of its word order domain can appear discontinuously in the mother's domain as long as the relative order of the elements is preserved in the mother's domain. Here too, the building up of word order domains interacts with recursion, which means that the size of word order domains may not be bounded by a given grammar.

3.2 Application

We assume an architecture of the prosody-syntax interface along the lines of Dalrymple & Mycock (2011). That is, we assume that the grammar builds syntactic and prosodic trees in tandem and that trees meet in the s- and p-strings, which are associated via their co-occurrence in lexical entries: the string is therefore the sole point of interface between syntax and prosody. Following Mycock & Lowe (2013), we do not use dedicated projections (e- and χ-structure) to pass information through the prosodic and syntactic trees, but rather assume that the terminal elements of the s- and p-strings are AVMs that can store information beyond the form of the relevant string elements.

Concretely, we need the p-string to contain information about prosodic hosting patterns. This is done via annotation on the prosodic structure building rules in the same way as in Mycock & Lowe (2013). As discussed in section 2.1.1 there are two patterns that are special to clitics: recursive prosodic word formation (example 4 above) and hosting of the clitic by a subprosodic word (example 7 above). The latter process is less well understood, but for concreteness we will assume stray adjunction of syllables. The relevant prosodic structure building rules are given in (18)–(19).[7]

$$
\begin{array}{llccc}
(18) & \omega & \to & \omega & \sigma \\
& & & \omega \in (\nearrow \mathrm{L}) & \omega \in (\searrow \mathrm{R}) \\
& & & \mathrm{IntP} \in_c (\nearrow \mathrm{L}) & (\searrow \textit{HOST}) = \omega
\end{array}
$$

$$
\begin{array}{llcccc}
(19) & \omega & \to & \sigma & \sigma^* & \sigma \\
& & & \omega \in (\nearrow \mathrm{L}) & & \omega \in (\searrow \mathrm{R}) \\
& & & \mathrm{IntP} \in_c (\nearrow \mathrm{L}) & & (\searrow \textit{HOST}) = \sigma
\end{array}
$$

---

[7]Following Mycock & Lowe (2013), we use italics for p-string features.

The first line in each of these rules simply passes down edge information to the terminal left ($\swarrow$) and right ($\searrow$) daughters.[8] The interesting things happen in the second line: ($\downarrow$ *HOST*) $= \omega|\sigma$ records the type of prosodic host (syllable or prosodic word) in the p-string of the clitic, and IntP $\in_c$ ($\swarrow$ L) ensures that hosting can only apply at the left edge of IntP. Notice that the rules leave open whether the relevant IntP is the core clause or a verse-induced IntP as in (12).

We now come to the lexical entry of clitics. As we saw in section 2.1.1, sentential clitics are typically happy to accept both a PW and a sub-PW as their host, whereas clausal clitics require a PW host. (20) gives sample lexical entries for the sentential clitic *gár* (roughly 'for', signalling a causal connection between the sentence in which it occurs and some piece of preceding discourse) and the clausal clitic *me* ('me').

(20)   *gár* $\widehat{\mathrm{Adv}}$ ($\searrow$ *HOST*)   *me* $\widehat{\mathrm{D}}$ ($\searrow$ *HOST*) $=_c \omega$

$\qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad\qquad\qquad$ ($\uparrow$ PRED) = 'PRO'

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ ($\uparrow$ CASE) = ACC

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots$

The constraints on the *HOST* feature ensure that *me* can only be inserted in the prosodic structure via (18) whereas *gár* can be inserted with (18) or (19). To keep yield functions from overgenerating, we assume that the yield functions h, r and i modify the productions they apply to by inducing a syntactic existential constraint ($\uparrow$ HOST) on the argument hosted by that yield function (i.e., the $n-1$ first daughters for h_n, r_n and the $n+1$'th daughter for i_n), with the effect that non-concatenative yield functions on the syntactic side must be licensed by (18)–(19) in the prosody.[9]

This is in fact all we need to derive clitic behaviour. On the prosodic side, the *HOST* feature requires the clitic to go after an appropriate host at the left edge of its IntP. Since the prosodic and syntactic structures are built over the same string, with no reordering, the syntax must position the clitic appropriately. Notice that no information passing between prosody and syntax is required; prosody influences syntactic positioning because both structures must be simultaneously well-formed.

To see how this works, consider (10). The core clause S in this example is *pánta kakà ékʰein*, which by our assumptions forms an IntP. So the clitic *spʰi* must find a suitable host at the left edge of this IntP while at the same time being a daughter of S. Since in this case the first constituent of S, *pánta*, is exactly one prosodic word and the clitic qua $\widehat{\mathrm{D}}$ can be the second constituent of S, this is straightforward and requires no yield functions beyond concatenation. For the sentential clitic *gàr* things are somewhat more complicated: it must find a prosodic host within the

---

[8]As noted by Mycock & Lowe (2013), these can probably be stated as more general principles and need not actually be stated on every rule.

[9]This requires a slight change to the principle of Interface Harmony (Dalrymple & Mycock, 2011; Mycock & Lowe, 2013) since an existential constraint on the syntactic side is verified by a constructive constraint on the prosodic side.
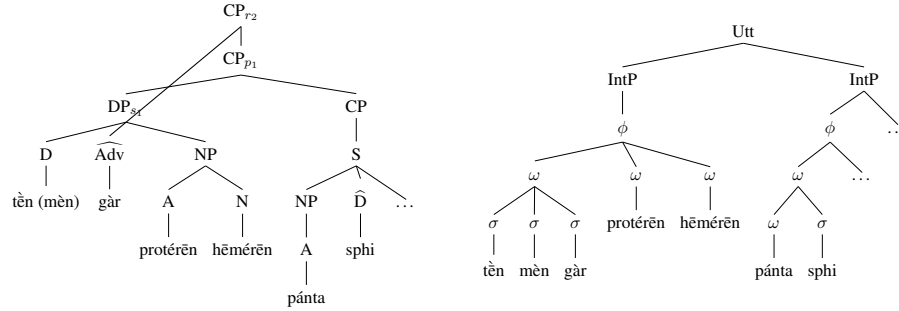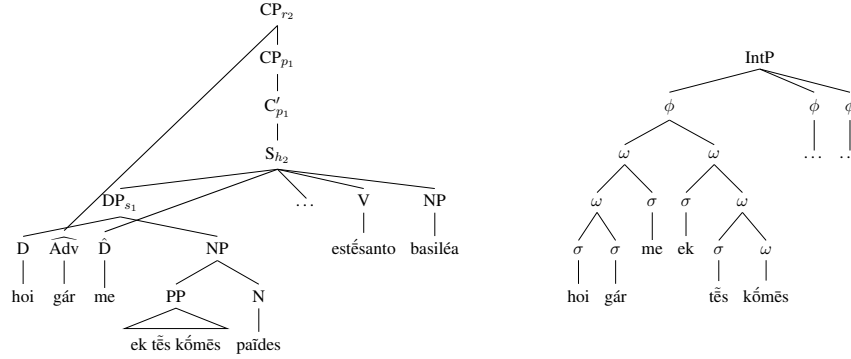
Figure 2: Syntactic and prosodic structure for (10)

Figure 3: Syntactic and prosodic structure for (8i)

topicalized phrase *tèːn=mèn protéreːn heːmέreːn* which has been adjoined to CP and forms a separate IntP. At the same time it must adjoin to the highest CP, i.e. above the phrase which must host it prosodically. This can be done by exploiting the yield functions p_1 and r_2, as shown in Figure 2.[10]

We can also derive the correct behavior in sentences like (8i) where there is no S-external material apart from the sentential clitic. Such sentences make up only one IntP and consequently, the sentential clitic must be hosted lower than its own position in the syntactic structure. This is achieved by hosting the clausal clitic in a gap in the first constituent DP in S, and then propagating this gap up to the CP-level, where it is resolved, as shown in Figure 3.

A salient feature here is that the host D *hoi* is not itself a prosodic word, but forms one with the first clitic, which can take a syllable host. The result is a prosodic word, which can therefore host *me*. Finally, we note that we can capture the complex example in (11ii) where the clitic is hosted in Spec,CP even if it belongs in the S domain. This is shown in Figure 4.[11]

---

[10]Needless to say, given our knowledge of AG, much of the prosodic structure assumed in the following figures is based on conjecture. But all the points crucial to our analysis have been argued for here.

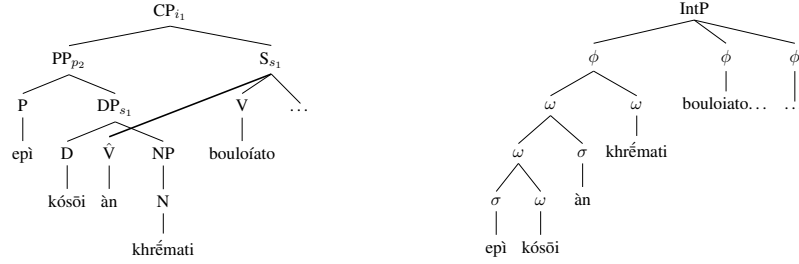[11]We add that there are also examples where Spec,CP is apparently not in the same IntP as S even

Figure 4: Syntactic and prosodic structure for (11ii)

4 Discussion

The main advantage of our proposal, we contend, is that we can treat clitics as syntactically normal words as far as the categorial structure goes.[12] What is special about them is their need for a prosodic host, which drives their special linearization. MCFG is the perfect tool to capture this, with its distinction between categorial structure and linearization (i.e., yield computation).

By contrast, all other approaches try to put clitics in the "correct" position in an ordinary, context-free c-structure, where they can get the correct grammatical function, and then displace them at some point in LFG's projection architecture. For Bögel et al. (2009, 2010) and Bögel (2015), they are displaced between the syntactic and the prosodic string, i.e. we have a syntax-prosody mismatch. For Lowe (2015), they are displaced between the c-structure and the syntactic string, i.e. we have a syntax-internal mismatch where the yield of the c-structure is not identical to the syntactic string. Architectural differences apart, there are similar conceptual and empirical problems with both approaches: both need to motivate a non-surface position in the c-structure and a means of enforcing this position.

4.1 Non-surface positions

Traditionally, it has been a hard constraint on LFG c-structures that their yield should match the output string. This means that the precedence order on the string can guide our assumptions about the precedence order on the terminals of the c-structure. If we give up this assumption, we need other heuristics to determine c-structure position.

Consider our examples (8i) and (10) with Figures 2–3. On a displacement view, we must motivate non-surface positions for the clitics in these trees. For *gár* this is perhaps not too difficult, since this particle has sentential scope and so we can assume that it adjoins high up and hence its underlying position is to the left of

---

in the absence of a focalized constituent and where consequently clausal clitics are hosted after the first prosodic word in S, discounting Spec,CP. We have as of yet no account for this variation.

[12]Apart from their non-projecting status, that is, but non-projection (to the extent it is needed) is independently motivated for these words and in fact not essential to derive their clitic behavior.

the rest of the sentence. But consider *me* in Figure 3. On the displacement view, this must be realized to the left of the rest of S in order to be targeted by prosodic inversion. But what is the evidence that *me* actually has this c-structure position? None of the traditional LFG heuristics work: AG does not have a dedicated object position, and the stressed variant *emé* can occur in various positions in the clause, both pre- and post-verbally. The assumption that *me* is S-initial is entirely driven by the need to collect all clitics in a single position where they can be targeted by prosodic inversion. Even worse, the displacement view is forced to assume that *sphi* in Figure 2, which satisfies all the prosodic and syntactic constraints in its surface position, is actually displaced from an underlying S-initial position. Again, there seems to be no motivation for this assumption, except the need for the clitic to undergo inversion. Finally, it is unclear how to derive the position of *an* in Figure 4 on a prosodic inversion view: for functional reasons *an* belongs inside S, but it actually surfaces to the *left* of the edge of S.

4.2   Clitic as a syntactic category

While the assumption that clitics have a dedicated covert position in the c-structure is doubtful, the means of enforcing this position is no less problematic. Theories of this kind standardly assume that all clitics belong to a single lexical category CL. But clisis is a prosodic category, not a functional one. Cross-linguistically and even within single languages such as AG the functional categories of clitics are so diverse (encompassing at least pronominal elements, connectives, discourse particles, and tense and modal auxiliaries) that a single unified category is unappealing (O'Connor, 2002, 316).

Moreover, sequences of clitics are typically hosted under a CCL node (e.g., examples 16–17 in Lowe 2011). There is, however, little to no evidence that clitic clusters form a syntactic or a prosodic constituent. This analysis also assigns clitics with diverse functional profiles (pronouns, auxiliary verbs, discourse particles, modal quantifiers) the same lexical category. The lexical category is in fact defined by the need for a prosodic host. We see little appeal in this move, because lexical categories and prosodic properties are independent dimensions of lexemes. By allowing them to interact, one opens the door to a lexical category for all stressed words, which is an unwelcome possibility. In contrast, our approach lets clitics have exactly the category that we expect from their functional profile.

4.3   Comparison with linearization HPSG

As we saw in section 3.1, yield functions as found in MCFG are a less powerful device than the word order domains found in linearization-based HPSG, which were developed to distinguish between tectogrammatical and phenogrammatical structure, something which is already present in the c-/f-structure distinction in LFG. Since linearization is a more powerful device, designed to do more than what our yield functions do, it must be constrained in order not to overgenerate. For

example, the analysis of Penn (1999) requires four word order domains, the pre-clitic, clitic, post-clitic and remainder fields, in order to prevent post-clitic material that belongs syntactically with the pre-clitic field to mix with the remainder field. In our approach, there is never more than a binary split of the yield.

But modulo the formal device involved, the basic idea is the same in our approach and that of Penn (1999) and Crysmann (2006). In the HPSG version, prosody can effect compaction, whereas in our approach non-standard yield functions are applied for essentially prosodic reasons. In HPSG, prosodic and syntactic compaction is modelled in the same structure, which predicts a much closer relationship between syntax and prosody than we assume. That, however, seems to be a result of the underdeveloped role of prosody in HPSG: there is only one level where constraints can be satisfied, but as Penn (1999, 6) points out, "it is entirely unsatisfying to represent prosodic structure as a disconnected list of prosodic words to be carried around in an otherwise syntactic derivation. Ultimately, the notion of sign in HPSG must be changed to allow parallel derivations of prosody, syntax and discourse." Our LFG account gets this for free on account of the parallel architecture of the framework.

4.4   Architectural issues

Increasing the generative power of the c-structure from a CFG to a 2-MCFG does not increase the complexity of the combined LFG formalism. As shown by Seki et al. (1993), any MCFG can be translated into an LFG, and any LFG which bounds the number of c-structure nodes corresponding to each f-structure can be translated into an MCFG. Though we omit the proof, the same is true for the MCFG c-structure component of our analyses: the linearization facts that we analyze with the help of yield functions could in principle be captured with reentrancies, i.e. with multiple c-structure nodes mapping to a single f-structure. This would entail treating clitic-induced discontinuities in the same way as scrambling, but it would require a lot of linguistically unmotivated bookkeeping in the f-structure to enforce the significant differences between clitic-induced discontinuities and scrambling, which for example can never separate a determiner from its noun phrase.

Similar problems could arise if one wanted to apply our MCFG-approach to e.g., scrambling discontinuities. At present it is not clear to us what the domain of application for non-concatenative yield functions is. We have assumed here that they are governed by the HOST-feature, which means that only clitics can have non-concatenative syntax, but future research may show that the domain of application is wider. A larger question is whether grammars can vary on this point or whether there are universal principles of non-concatenative syntax.

Finally, one aspect that does not fall out from our account, but which is stressed by Bögel et al. (2010), is the finiteness of the set of admissible clitic sequences in any given language. This is predicted by their account because it models the syntax-prosody relation as a regular transducer, which can only handle a finite set of elements to be displaced at the interface. There is nothing directly comparable

in our theory, but it may be possible to prove that finiteness of the set of possible fillers of gaps in two-dimensional yields decreases the parsing complexity of the MCFG grammar, perhaps down to that of an ordinary CFG.

5 Conclusion

We have argued that the distribution of second-position clitics motivates an increase in the power of c-structure to that of a 2-MCFG in analyzing natural language syntax. Our analysis obviates both covert positions in c-structure and the assumption that "clitic" is a syntactic category. Adopting an MCFG-based c-structure does not increase the expressivity of the LFG formalism, but it does mean that c-structure can describe more complex phenomena that would otherwise be left to other projections in the LFG architecture. We leave for future research the question of whether the yield functions of MCFG could be used to insightfully model other linguistic phenomena.

References

Agbayani, Brian & Chris Golston. 2010a. Phonological movement in classical Greek. *Language* 86. 133–167.

Agbayani, Brian & Chris Golston. 2010b. Second-position is first-position. *Indogermanische Forschungen* 115. 1–21.

Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. Oxford: Oxford University Press.

Anderson, Stephen R. 2012. Clitics. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 4, 2002–2018. Malden, MA: Blackwell.

Bögel, Tina. 2015. *The syntax–prosody interface in Lexical Functional Grammar*: Universität Konstanz dissertation.

Bögel, Tina, Miriam Butt, Ronald M. Kaplan, Tracy Holloway King & John T. Maxwell III. 2009. Prosodic phonology in LFG. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG09 conference*, 146–166. Stanford: Center for the Study of Language and Information.

Bögel, Tina, Miriam Butt, Ronald M Kaplan, Tracy Holloway King & John T Maxwell. 2010. Second position and the prosody-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Fifteenth international lexical functional grammar conference (LFG10)*, 106–126. Stanford: Center for the Study of Language and Information.

Ćavar, Damir & Melanie Seiss. 2011. Clitic placement, syntactic discontinuity, and information structure. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG11 conference*, 131–151. Stanford: Center for the Study of Language and Information.

Chandler, Henry W. 1881. *A practical introduction to Greek accentuation*. Oxford: Oxford University Press 2nd edn.

Clark, Alexander. 2014. An introduction to multiple context free grammars for linguists. http://www.cs.rhul.ac.uk/home/alexc/lot2012/mcfgsforlinguists.pdf.

Crysmann, Berthold. 2006. Floating affixes in Polish. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Varna*, 123–139. Stanford: Center for the Study of Language and Information.

Dalrymple, Mary & Louise Mycock. 2011. The prosody-semantics interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG11 conference*, 173–193. Stanford: Center for the Study of Language and Information.

Devine, Andrew M. & Laurence D. Stephens. 1994. *The prosody of Greek speech*. Oxford: Oxford University Press.

Devine, Andrew M. & Laurence D. Stephens. 1999. *Discontinuous syntax*. Oxford: Oxford University Press.

Dik, Helma. 1995. *Word order in ancient Greek*. Amsterdam: Gieben.

Fortson, Benjamin W. 2010. *Indo-European language and culture*. Malden, MA: Blackwell 2nd edn.

Fraenkel, Eduard. [1933] 1964. Kolon und Satz. In *Kleine Beiträge zur klassischen Philologie*, vol. 1, 93–130. Roma: Storia e Letteratura.

Goldstein, David. 2016. *Classical Greek syntax: Wackernagel's law in Herodotus*. Leiden: Brill.

Hale, Mark R. 1987a. Notes on Wackernagel's law in the language of the Rigveda. In Calvert Watkins (ed.), *Studies in memory of Warren Cowgill (1929–1985)*, 38–50. Berlin: de Gruyter.

Hale, Mark R. 1987b. *Studies in the comparative syntax of the oldest Indo-Iranian languages*: Harvard University dissertation.

Kathol, Andreas. 1995. *Linearization-based German syntax*: Ohio State University dissertation.

Kuhlmann, Marco. 2013. Mildly non-projective dependency grammar. *Computational Linguistics* 39. 355–387.

Lowe, John J. 2011. Ṛgvedic clitics and 'prosodic movement'. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG11 conference*, 360–380. Stanford: Center for the Study of Language and Information.

Lowe, John J. 2014. Accented clitics in the *Ṛgveda*. *Transactions of the Philological Society* 112. 5–43.

Lowe, John J. 2015. Clitics. *Journal of Linguistics* 52. 375–419.

Mycock, Louise & John Lowe. 2013. The prosodic marking of discourse functions. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG13 conference*, 440–460. Stanford: Center for the Study of Language and Information.

O'Connor, Robert. 2002. Clitics and phrasal affixation in constructive morphology. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG02 conference*, 315–332. Stanford: Center for the Study of Language and Information.

Penn, Gerald. 1999. A generalized-domain-based approach to serbo-croatian second position clitic placement. In Gosse Bouma, Erhard W. Hinrichs, Geert-Jan M. Kruijff & Richard T. Oehrle (eds.), *Constraints and resources in natural language syntax and semantics*, 119–136. Stanford: Center for the Study of Language and Information.

Pollard, Carl. 1984. *Generalized phrase structure grammars, head grammars, and natural language*: Stanford University dissertation.

Probert, Philomen. 2003. *A new short guide to the accentuation of ancient Greek*. London: Bristol Classical Press.

Reape, Mike. 1994. Domain union and word order variation in german. In John Nerbonne, Klaus Netter & Carl Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, 151–197. Stanford: Center for the Study of Language and Information.

Seki, Hiroyuki, Ryuichi Nakanishi, Yuichi Kaji, Sachiko Ando & Tadao Kasami. 1993. Parallel multiple context-free grammars, finite-state translation systems, and polynomial-time recognizable subclasses of lexical-functional grammars. In *Proceedings of the 31st Annual Meeting on Association for Computational Linguistics (ACL '93)*, 130–139. Morristown, NJ: Association for Computational Linguistics. doi:10.3115/981574.981592.

Taylor, Ann. 1990. *Clitics and configurationality in ancient Greek*. Philadelphia: University of Pennsylvania dissertation.

Toivonen, Ida. 2003. *Non-projecting words*. Dordrecht: Kluwer.