



WestminsterResearch

<http://www.wmin.ac.uk/westminsterresearch>

In favor of a "clitic cluster" in the Bulgarian and Macedonian DP.

Andrew Caink *

School of Social Sciences, Humanities and Languages

* At the time of publication Andrew Caink was employed by the University of Wolverhampton

This is an electronic author version of a paper published in Dimitrova-Vulchanova, Mila, Dyer, Donald L., Krapova, Iliyana and Rudin, Catherine, (eds.) *Papers from the Third Conference on Formal Approaches to South Slavic and Balkan Languages*. Balkanistica, Issue 15, Mississippi: University of Mississippi Printing Services, pp. 75-102.

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners. Users are permitted to download and/or print one copy for non-commercial private study or research. Further distribution and any use of material from within this archive for profit-making enterprises or for commercial gain is strictly forbidden.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch.
(<http://www.wmin.ac.uk/westminsterresearch>).

In case of abuse or copyright appearing without permission e-mail wattsn@wmin.ac.uk.

In favour of a ‘Clitic Cluster’ in the Bulgarian and Macedonian DP¹

Andrew Caink

University of Wolverhampton

1. Introduction

In this paper, I argue for the existence of a ‘clitic cluster’ in the Bulgarian and Macedonian DP in (1), on a par with the clitic cluster in the clause.

- | | | |
|--|--|-------------|
| (1)a. <i>Kniga -ta -mi</i>
book ·FEM.SG· the ·FEM.SG· my ·DAT·
‘My book/that book of mine’ | b. <i>Goljemi-te -mi drexi</i>
big ·PL· the ·PL· 1sg ·DAT· clothes ·PL·
‘My big clothes’ | (Bulgarian) |
|--|--|-------------|

This clitic cluster maximally consists of the determiner and the dative clitic in the strict order shown above. In Caink (1997, 1999), the term ‘clitic cluster’ is not an informal observation but a primitive of the system. This paper outlines the treatment of South Slavic closed class items argued for there and demonstrates the role played by the cluster in the inter-related distribution of the determiner and dative clitic.

I demonstrate that the distribution of the post-posed determiner and (Macedonian) demonstratives provide confirming evidence for the ‘Invisible Category Principle’ (Emonds 1987), whereby a closed class feature associated with one node in a tree may be alternatively realized via sisterhood on another node. Whilst assuming a theoretically similar account of the dative clitic, I show that the placement of the dative clitic requires a revised Alternative Realization argued for in Caink (1998) in an analysis of the clausal clitic cluster in Serbian/Croatian/Bosnian.

Having established the nature of the determiner and dative clitic and the mechanism by which they appear, I argue that their interrelated distribution can only be accounted for if they are lexicalized together at PF following construction of the clitic cluster in the numeration. This account is founded upon a unified theory of syntactic categories argued for in Emonds (1985) and subsequent works, and makes use of independently motivated mechanisms. The analysis avoids the often *ad hoc* movement rules that are a characteristic of alternative generative accounts of this data (e.g. Franks 1998, Arnaudova 1996, Dimitrova-Vulchanova & Giusti 1999) and provides theoretical definition to the term ‘clitic cluster’ that is generally lacking in formal accounts.

Section 2 reviews the basic data for the determiner/demonstratives and the possessive dative clitic and some of the difficulties encountered in the literature in coping with such data. I establish two descriptive generalisations to be addressed in this article. Section 3 introduces the original formulation of the ‘Invisible Category Principle’ (ICP) and how it predicts the distribution of the post-posed determiner/demonstrative in Bulgarian and Macedonian. Section 4 introduces the problem of the ICP in accounting for the dative clitic and provides a revised definition in terms of extended projections rather than sisterhood. Section 5 discusses the nature of the clitic cluster and the role it plays in the distribution of the post-posed determiner/demonstrative and dative clitic. Section 6 presents a summary.

2. The Data and Descriptive Generalisations

In this section, I introduce the basic data for the determiner/demonstratives and dative clitics in turn and consider some of the difficulties encountered in generative movement-based analyses. I will then establish the descriptive generalisations to be addressed in subsequent sections.

2.1. Post-posed determiners and demonstratives: the data

The table in (2) displays the person and number paradigm for the determiners in Bulgarian and Macedonian.

¹ I am grateful to Ruslan Mitkov, Violeta Sotirova and Olga Arnaudova for help with data and to Steven Franks and Zeljko Bošković for comments on the original paper. Any errors are mine.

(2) *The Bulgarian and Macedonian determiners*

	Singular	Plural
Masculine	<i>Bulgarian</i> <i>Macedonian</i>	-â(t) / -ja(t) -ot
feminine	-ta	-te/-ta
neuter	-to	

Basic examples in the DP are shown in (1) above and (3) here:

- (3) a. *Golemi -jât mi blok* (Bulgarian)
 big def._{MASC.SG} 1sg._{DAT} block_{MASC.SG}
 'My big block'

b. *Visok -iot čovek* (Macedonian)
 tall def._{MASC.SG} man
 'The tall man'

These preliminary data in (1) and (3) illustrate the fact that the determiner does not always follow the head noun of the DP; if the noun is modified by an adjective as in (1b)/(3a), then the determiner follows the adjective. We shall return to this distribution below.

In Bulgarian, demonstratives are realized by free morphemes (i.e. not ‘bound’) (see (8) below), but Macedonian demonstratives are realized as post-position bound morphemes in the same way as determiners:

- (4) *Kniga -ya/na* (Macedonian)
book this/that (Franks 1998:60)

The paradigm in (2) is a simplification: it is well-documented in the literature that for some nouns, the form of the determiner is not decided by person/number but on phonological grounds alone (e.g. Scatton 1984; Franks 1998:61). Consider (5):

- (5) a. *čičo* -*to* (Bulgarian)
 uncle._{MASC.SG} -the._{NEUT.SG}
 b. *bašta* -*ta*
 father._{MASC.SG} -the._{FEM.SG}

In (5a), the masculine singular noun *čičo* ‘uncle’ ends in *-o* and selects the determiner *-to* in the same way as neuter nouns ending in *-o* (*pismo -to* ‘the letter’). In (5b), the masculine noun selects an apparently feminine determiner *-ta* on account of its ending with the vowel *-a* like canonical feminine singular nouns (*kniga -ta* ‘the book’). However, phonology does not always over-ride the morphology of number/person:

- (6) a. *Sto -te sela* (Bulgarian)
 hundred the villages._{NEUT.PL}
 'The hundred villages'

b. *Mnogo -to novi knigi*
 many the new._{FEM.PL} books._{FEM.PL}
 'The many new books'

In (6a), the word-final vowel on *sto* ‘hundred’ usually takes the neuter singular determiner *-to* (e.g. *pismo -to* ‘the letter’, or *čičo -to* ‘the uncle’ in (5a)), yet the canonical plural form *-te* is selected. In contrast, another plural modifier in (b), *mnogo* ‘many’, appears to follow phonological information alone, despite the related semantics.

In addition to such cases, there are many instances in Bulgarian where alternative plural forms of nouns exist (e.g. Scatton 1984:425-6): for example *ramo* 'shoulder', *ramene ~ ramena* 'shoulders'.

These two plural forms select distinct determiners: *ramene -te* / *ramena -ta* ‘the shoulders’. Such facts must be listed in the lexicon.

The conclusion that Franks (1998:61) comes to is that the required determiner is selected sometimes idiosyncratically (hence must be stipulated in the lexicon), sometimes it results from morphological information for person/number, and sometimes it is decided on purely phonological grounds. In other words, it resembles the behaviour of *inflectional morphology* such as the bound person/number morphemes on finite verbs.

Turning now to the distribution of the determiners and demonstratives², we have already seen in (1a) that when only a noun is present, the determiner is suffixed to the noun. If the noun is modified by an adjective, the determiner is suffixed to the adjective, as in (1b) and (3). If more than one adjective modifies the noun, the determiner appears on the first adjective, as in (6b) above.

In (7), it is clear that the determiner suffixes to the *head* of the first modifying AP:

- (7) a. *Ne samo* [_{AP}[*izklučitelno umno -to*] *kuče*]
not only extremely smart the dog
'Not only the extremely smart dog'
- b. **Ne samo* [_{AP}[*izklučitelno -to umno*] *kuče*]
not only extremely the smart dog
- c. [_{AP}[*Dosta glupava-ta*] *kniga*]
quite stupid the book
'The quite stupid book'
- d. *[_{AP}[*Dosta -ta glupava*] *kniga*]
quite def. stupid book

In (7b,d), the determiner cannot be suffixed to the specAP *izklučitelno* ‘extremely’ or *dosta* ‘quite’ respectively. This indicates that syntactic structure is a factor in determiner placement: it is not a simple ‘second position’ phenomenon that can be adequately dealt with in the phonology.

Let us now turn to consider a more articulated DP that includes a demonstrative. In traditional standard Bulgarian, the bound determiner is said to appear in complementary distribution with the demonstratives *tozi/tazi/tova* ‘this’ and *onzi/onazi/onova* ‘that’.

- | | |
|--|-----------------------------|
| (8)a. <i>Tazi kniga</i> / <i>kniga -ta</i> | b. * <i>Tazi kniga -ta</i> |
| this book book the | |
| c. <i>Onovo pismo</i> / <i>onovo -to</i> | d. * <i>Onovo pismo -to</i> |
| that letter letter the | |

However, Arnaudova (1998) points out that colloquial Bulgarian regularly allows the demonstrative and determiner to co-occur.

- | | |
|---|---|
| (9)a. <i>Tazi kniga -ta</i> | b. <i>Tazi hubava -ta kniga</i> |
| this. _{FEM} book. _{FEM} the. _{FEM} | this. _{FEM} nice. _{FEM} the. _{FEM} book. _{FEM} |
| 'This book' | 'This nice book' |
| c. <i>Tozi čovek -ja</i> | d. <i>Tozi neprijatni -ja čovek</i> |
| this. _{MASC} man. _{MASC} the. _{MASC} | this. _{MASC} unpleasant. _{MASC} the. _{MASC} man. _{MASC} |
| 'This man' | 'This unpleasant man' |
| e. <i>Tozi sčasti *(-ja) čovek</i> | f. <i>Tova tsjalo *(-to) čakane</i> |
| this. _{MASC} same. _{MASC} the. _{MASC} man. _{MASC} | that. _{NEUT} whole. _{NEUT} the. _{NEUT} waiting. _{NEUT} |
| 'The very man' | 'This whole waiting' |
| g. <i>Vsicki -te tezi knigi</i> | |
| all. _{PL} the. _{PL} these. _{PL} books. _{FEM.PL} | |
| 'All these books' | |
- (Arnaudova 1998:7-8 and pers. comm.)

² For the remainder of this article we shall focus on Bulgarian examples.

In each example in (9), a demonstrative co-occurs with the underlined determiner. Examples (9e,f) are in fact standard and become ungrammatical *without* the determiner, as shown.

Nothing in the following account turns on whether co-occurrence of the demonstrative and determiner is possible in Bulgarian but, significantly, both possibilities are allowed for in this theory. Next I review the sort of difficulties that are encountered in movement accounts of the determiner.

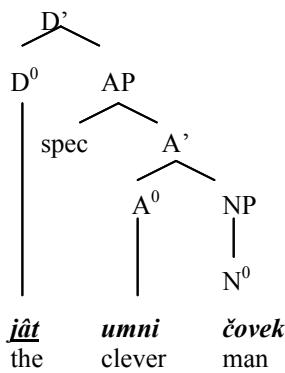
2.2. Difficulties for movement-based accounts of the determiner/demonstrative

The central problem in a purely movement-based account of the determiner (e.g. Fowler & Franks 1994, Arnaudova 1996, Franks 1998, Dimitrova-Vulchanova & Giusti 1999) is how the determiner ends up suffixed to the S first head down from the DP. This remains true whatever phrase structure for the DP one adopts, either that of (10b) in which a modifying AP forms part of the extended projection of the N, or that of (10c) where a modifying AP appears in the specifier position of a functional head (such as in the general framework of Cinque 1997).

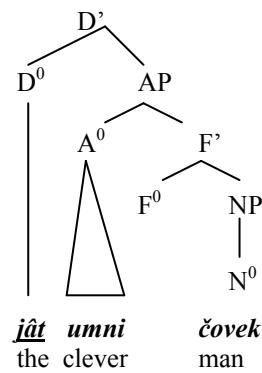
- (10)a. *Umni ját čovek*

clever the man
'The clever man'

b.



c.



Assuming the determiner is generated in D^0 , in (10b) it has to move down to suffix to the A^0 *umni* 'clever'; or in (10c), the determiner must move down to the head of the specifier of its complement. In both cases, downwards movement is problematic in a Principles & Parameters framework because it results in the foot of the chain c-commanding its head.

Alternatively, if one argues that the A^0 in (10b) raises to a position higher than D^0 , we run into the problem of explaining data such as (7) in which a full AP clearly precedes the determiner. The only way the AP minus its complement NP can raise is if the complement of the A^0 scrambles out of the AP first, leaving only a remnant AP to raise. This would be obligatory each time the DP includes an adjectival modifier. Worse still, in a case such as (6b) where the DP includes more than one modifying AP, it is all but the highest AP that must scramble. In the absence of any independent motivation for such scrambling, such an account remains grossly *ad hoc*.

Remaining with such a raising analysis, if one adopts the structure in (10c), the question arises as to what the motivation of AP raising *is*, a question of crucial importance in Minimalist approaches (Chomsky 1995). To argue that the AP in (10c) raises because of the prosodic requirements of the determiner (i.e. because it is marked 'suffix' in the lexicon) requires that such prosodic information is present in the syntax (what is often informally called 'look ahead'). This involves a very different architecture to that employed in Principles & Parameters models, not least requiring a bi-directional interface between the syntax and phonology (though see Zec & Inkelaas (1990) for such a model in their discussion of Serbian/Croatian clitics). Instead, if one argues that an abstract feature motivates any movement to a pre-determiner, it is difficult to see what kind of feature it is that attracts either an NP or AP.

A further alternative might be to assume a strong lexicalist approach whereby the determiner is suffixed to the head prior to insertion into the syntactic component. Raising is then for the purposes of checking the [+DEF] feature in D^0 against the determiner morpheme. However, this would be unable to rule out the situation where the determiner is suffixed to the 'wrong' head, such as the noun in a DP with modifying AP: for example, rather than the correct *Novi-te knigi* 'the new books', we would be

unable to rule out $*_{NP}[Knigi-te]_i novi t_i$, where the determiner is suffixed to the noun and the NP has raised above the AP headed by *novi* ‘new’ to check the [+DEF] feature in DP.

To account for the determiner placement via movement in the phonology is even more problematic, given that there is no independent evidence for phonological movement, and simple mechanisms that switch the linear order of two items (such as employed in Prosodic Inversion accounts of the Wackernagel Position, Halpern 1995) are too weak here: the data in (7) rules out a simple linear switch because it is the syntactic information ‘head’ that is central to the placement of the determiner.

The descriptive generalisation that gives rise to the difficulties discussed above is that the determiner appears suffixed to the head of the complement to D^0 , whether that complement be the first of several modifying APs or the bare NP, as established in Franks (1998).

- (11) **Descriptive generalisation #1:** *the determiner is an inflection which appears on the (head of the complement to D^0)* (Franks 1998)

As we have seen above, this causes a number of problems for movement based accounts of the determiner. However, we shall see in section three that (11) is exactly predicted by the independently motivated Invisible Category Principle of Emonds (1987) for purely closed class morphemes. First however, I review the placement of the other member of the DP clitic cluster, the possessive dative clitic.

2.3. *The possessive dative clitic: the data*

The paradigm for the possessive dative clitic is shown in (12).

- (12) Bulgarian Possessive clitics

1sg.	<i>mi</i>	1pl.	<i>ni</i>
2sg.	<i>ti</i>	2pl.	<i>vi</i>
3sg.masc./neut.	<i>mu</i>	3pl.	<i>im</i>
3sg.fem.	<i>i</i>	reflexive	<i>si</i>

The possessive clitic appears adjacent to any determiner in the DP:

- (13) a. ***Himikal-ât-mi***
pen the 1sg._{DAT}
'the pen of mine'
- b. ****Himikal -mi***
pen 1sg._{DAT}
'A pen of mine'
- c. ****Kupih edna -mu nova kniga***
bought-1sg. one 3sg._{DAT} new book
'I bought one new book of his'
- d. ***Haresvam nova -ta-mu šapka***
Like-1sg. new the 3sg._{DAT} hat
'I like his new hat'
- e. ****Haresvam nova -tašapka -mu***
Like-1sg. new the hat 3sg._{DAT}

Examples (13b,c) demonstrate that the possessive clitic cannot appear in the absence of a [+DEF] determiner³; (13d,e) indicate the strict adjacency requirement that holds between the determiner and the possessive clitic.

I argue that together, the determiner and possessive clitic form a clitic cluster in the same way as the pronominal clitics and clitic auxiliaries in the clause (see Caink 1998), with the maxim membership and relative order shown in (14a).

³ However, see (16) below in which the possessive dative appears adjacent to a lexical demonstrative.

- (14) a. *The Bulgarian/Macedonian DP clitic cluster:*
 {definite article/demonstrative} -- possessive clitic

- b. DP[AP[A[Dulgo] to mi NP[pismo]]]
 long.NEUT the.NEUT 1sg.DAT letter.NEUT
 'My long letter'

The example in (14b) illustrates the structure of a full DP containing the clitic cluster, which is enclitic on the first head of the complement to D⁰, *dulgo* 'long', as predicted by the generalisation in (11) for determiner distribution.

In one of the earliest generative treatments of Bulgarian clitics, Ewen (1979) establishes that there is a tight relationship between the distribution of the determiner and the possessive clitic. For example, some lexical items such as *vsjako* 'every' and *njakoj* 'someone' are in complementary distribution with the definite article, and are *also* in complementary distribution with the possessive clitic:

- (15) a. *Vsjako moe kuče e bolno*
 every my dog be-3sg. ill
 'Each one of my dogs is ill'
- b. **Vsjakoto moe kuče e bolno*
 every the my dog be-3sg. ill
- c. **Vsjako mi kuče e bolno*
- d. *Njakoj moj kolegi ne običat da karatski*
 Someone my colleagues not love to drive ski
 'One of my colleagues doesn't like to ski'
- e. **Njakoj-tomoj kolegi ne običat da karatski*
 Someone the my colleagues not love to drive ski
- f. **Njakoj mi kolegi ne običat da karat ski* (Ewen 1979:172)

In (15b,e), the determiner is incompatible with *vsjako* 'every' and *njakoj* 'someone', and (15c,f) indicates the same is true of the possessive dative clitic *mi* 'my'. The grammaticality of (15a,d) demonstrates that the restrictions in (15c,f) do not stem from a semantic incompatibility. This data underlines the fact that there is a close relation between the determiner and the possessive clitic, a relation that previous generative accounts have failed to capture satisfactorily. In this account, it is the existence of the clitic cluster and its lexicalization that ensures the strict adjacency between the determiner and the possessive dative clitic.

Next, note the significant fact that the possessive clitic can appear in one construction that lacks the determiner, those in which a demonstrative appears. Importantly, in this case, the possessive clitic appears adjacent to the [+DEF] demonstrative, which I assume is in D⁰⁴:

- | | |
|---|---|
| <p>(16) a. <i>Tezi ti knigi</i>
 these 2sg.DAT books
 'These books of yours'</p> <p>c. <i>Tozi mi himikal</i>
 this 1sg.DAT pen
 'this pen of mine'</p> | <p>b. <i>Tazi mu kola</i>
 this 3sg.DAT car
 'this car of his'</p> <p>d. <i>Tazi ni kušta</i>
 this 3pl.DAT home
 'This home of ours'</p> |
|---|---|

In each of the examples in (16), the possessive clitic follows the demonstrative. We can now formulate a descriptive generalisation that requires explanation in any formal account of the possessive clitic.

- (17) **Descriptive Generalisation #2:** *The possessive dative clitic appears adjacent to any determiner; if none is present, it appears adjacent to the lexical demonstrative.*

⁴ See Arnaudova (1998) for an alternative account of the demonstrative and its position in the DP.

We shall see in the next subsection that this generalisation proves very difficult to capture without a formal notion of a clitic cluster.

2.4. Difficulties in movement-based accounts of the possessive dative clitic

If movement-based accounts have a difficulty with accounting for the apparent downwards movement of the determiner or *ad hoc* raising of phrases within the DP, such approaches encounter even further difficulties in attempting to give an independently motivated analysis of the possessive clitic that captures the data in the previous subsection. Here we shall outline difficulties in the two approaches suggested in Franks (1998).

The first proposal assumes that the possessive clitic is the head of an Agreement Phrase, AgrIOP (Franks 1998:67). One benefit is that this establishes a connection with the dative pronominal clitics in the clause, which many have argued also head an AgrIOP (e.g. Rudin 1997).

- (18) DP[AGRIOP[AGRIOP[***mi***] AP[A[***nova-ta***] NP[[***kniga***]]] ‘My new book’
1SG.DAT new-the book

In (18), D^0 takes AgrIOP as a complement headed by the possessive clitic (for the moment, we set aside the problem of how the determiner appears suffixed to the adjective *nova* ‘new’). In this AgrIOP account, the clitic must *lower* to the next head. Given the absence of lowering movement in any version of the Principles & Parameters theory, this is simply a reformulation of the data. Indeed, it hardly achieves observational adequacy, given that the possessive clitic appears not to lower when D^0 is filled by a demonstrative as in (16). Again, as we saw in the case of the determiner, to stipulate that any movement occurs for prosodic reasons requires that the syntax somehow look ahead to PF in order to see which items need prosodic support and which do not.

Finally, the tree proposed in (18) nullifies the descriptive generalisation (11) in which it was observed that the determiner appears on the complement to D^0 . If we assume an Agreement Phrase in this way, we actually make a parsimonious account of the determiner even more distant.

The second alternative Franks (1998) suggests is to dispense with the notion that the possessive clitic projects its own syntactic phrase. Given that the clitic appears to be adjoined to the demonstrative in (16), he suggests that the clitic is base-generated adjoined to D^0 (though it remains unclear why) and if necessary, it lowers in the syntax to A^0 or N^0 . This retains the descriptive generalisation for the determiner (11) and avoids the wrinkle that agreement phrases are dispensed with in Chomsky (1995), but the analysis is still hamstrung by the need for *ad hoc* syntactic lowering and a lack of any discussion of the nature of this possessive clitic.

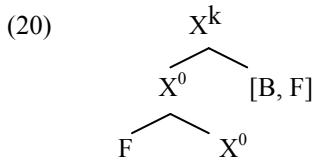
We have seen that the descriptive generalisations (11) and (17) prove difficult to capture, especially if they are treated as unrelated to each other. Fundamental to our account is the existence of a clitic cluster that binds these two morphemes in the numeration, hence the strict adjacency between them follows automatically.

3. The ‘Invisible Category Principle’ and the determiner ‘inflection’

Emonds (1987, 1994) argues that a given node in the syntax may remain ‘null’ (that is, it may dominate no phonological material) if the closed class formal features associated with that node are realised by a closed class morpheme elsewhere in the tree. A formal feature in this way is alternatively realized via a sisterhood relation between the original node and the morpheme that realises it:

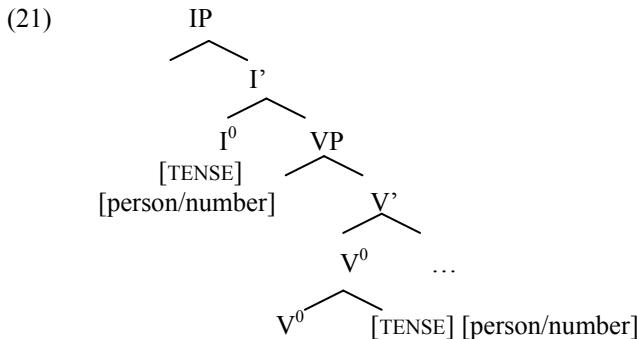
- (19) **Alternative Realization (AR):** A syntactic feature F matched in UG with category B can be realized in a grammatical morpheme under X^0 , provided X^k is a sister of [B, F].
(Emonds 1987)

To illustrate this, consider the tree in (20):



The formal feature F is associated with the node B in (20). This node may remain null if the feature F is alternatively realised on X^0 , X being in a sisterhood relation with B.

For example, consider the finite verb inflections in English which realize the features associated with I^0 by appearing on the verb. It has been well-established in the literature that the finite verb does not raise to I^0 to ‘pick up’ or ‘check’ the finite morphology before Spell-out in the way that finite verbs do in, say, French (Emonds 1978). In Emonds’ account, the inflectional morpheme is the alternative realization of the closed class features for Tense ([+/-PAST]), person and number (ϕ -features) associated with I^0 . This is illustrated in (21).

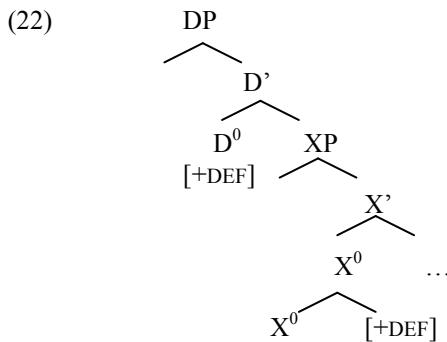


Note that (21) presents the simplest case; in fact, (19) allows for X^0 to constitute a ‘sister’ provided that *a projection of X* is a sister to B.

Other examples of this mechanism are the dative case ending in, say, Latin or the so-called semantic case endings in Finnish which in each case alternatively realize null P (see Nikanne 1993 for Finnish), or the English comparative and superlative variation between *more ~ -er* and *most ~ -est* respectively (Emonds 1987).

The distribution of the Bulgarian/Macedonian determiners and Macedonian demonstratives is clearly predicted by (19): these closed class ‘inflectional’ morphemes are the alternative realization of the [+DEF] feature associated with D^0 , appearing always on the head of a projection that is a sister to D^0 , as established in (11).

Interestingly, in the majority of cases that are cited as evidence for this principle in Emonds (1987, 1994, 1999), the host of a given morpheme is invariably the same for any given morpheme: tense and agreement morphemes in languages without V-movement to I always appear on a V, the feature associated with a closed class of P always appears on an N, arrived at via subcategorization frames for the given morpheme. However, (19) is couched in purely structural terms, hence makes no *necessary* restriction as to host. We should therefore expect to find cases where a given AR morpheme is more promiscuous in terms of its host yet still appears on a sister to the canonical node with which the formal feature is associated. This is precisely what happens in the case of the bound determiners and demonstratives in South Slavic:



The feature [+DEF] is always alternatively realized on X^0 , as we have seen. However, for independent reasons, X may be either noun [+N,-V] or an adjective [+N,+V]⁵.

We have seen then that the distribution of the South Slavic bound determiner and demonstrative morphemes are accounted for by the mechanism of Alternative Realization, or the ‘Invisible Category Principle’. Their distribution (and inflectional properties) are formally equivalent to the dative and ‘semantic’ case morphology on nouns and the inflectional morphology on finite verbs; no further theoretical innovation is required.

4. The ‘possessive dative clitic’ as an Alternative Realization of a null PP

In this section I address the nature and distribution of the possessive clitic. In the first part, a revision of the Alternative Realization mechanism is argued for and in the second part I outline the significant role played by subcategorization and phonological lexicalization.

4.1. A Revised Alternative Realization

In Caink (1998, 1999), it is argued that pronominal clitics in the South Slavic clause such as in (23) are further instances of Alternative Realization⁶.

(23)	<i>Mi</i>	<i>ja</i>	<i>dadoa</i>	<i>smetka</i>	<i>-ta_{PP}[Ø]</i>	(Macedonian)
	1SG.DAT	3SG.ACC	give-3PL.	bill	the	(Lunt 1952:38)

In (23), the pronominal clitic *mi* ‘to me’ is the alternative realization of the formal features [DATIVE] and [1st SG] on a null PP argument inside VP, shown as $_{PP}[\emptyset]$.

Before considering the second pronominal clitic *ja* ‘it’, recall that the formulation of Alternative Realization in (19) states that node B *may* be null if its formal features are alternatively realized. UG allows for this option, but also for the possibility that the node remains overt, in which case the AR morpheme is effectively ‘doubling’ the formal features of node B. One source, therefore, of language variation in those languages that employ the AR mechanism is whether the AR morpheme ‘doubles’ an overt node or not, and how a given language employs this mechanism. In Macedonian, ‘clitic doubling’ of [+DEF] direct object DPs is obligatory; in (23), the AR morpheme *ja* alternatively realizes the formal features [ACCUSATIVE] and [3rd SG] of the DP *smetkata* ‘the bill’, and this DP remains overt.

In fact, within this account, we have already seen a case of ‘doubling’ with respect to the determiner. In standard Bulgarian, the AR of the [+DEF] feature allows D^0 itself to be covert, but data in (9) showed that in colloquial Bulgarian and certain constructions with quantifiers, the AR morpheme doubles a lexical demonstrative in D^0 .

I will assume the same analysis for the possessive dative clitic found in the DP in South Slavic. One way of expressing possession in Bulgarian/Macedonian is via a PP headed by the functional preposition *na* ‘of/to’ which follows the possessed noun⁷. The dative clitic in the DP licenses this PP to be null, shown in (24), by alternatively realizing the formal features [DATIVE] and [3rd SG].

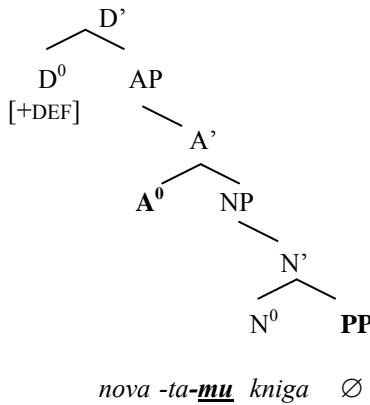
(24)a.	<i>Nova-ta mu</i>	<i>NP[kniga</i>	<i>PP[na-Ivan]]</i>	(Bulgarian)
	new -the	3SG.DAT book	of Ivan	

⁵ Alternatively, if we assume that a modifying AP appears in the specifier of a functional head in the extended projection of N, as in (10c), then X is a functional head. Either way, the determiner appears in a position immediately following the head of AP.

⁶ See also Emonds (1999) on Romance pronominal clitics.

⁷ Note that Ewen (1979) adopts Emonds’ typology of lexical categories (later published in Emonds 1985), and argues that Bulgarian *na* ‘of/to’ is subject to late lexical insertion, discussed in 4.2 below.

b.



nova -ta-mu kniga Ø

As predicted by (19), the tree in (24b) shows the determiner morpheme *-ta* ‘the’ on A^0 (highlighted), which is the head of the AP complement to D^0 . The possessive clitic also appears under A^0 , adjacent to the determiner.

However, whilst the position of the determiner in (24b) is predicted by Emonds’ formulation of Alternative Realization (19), the position of the possessive clitic is not: there is clearly *no* sisterhood relation between any projection of A^0 and PP. Indeed, the structural distance between the possessive clitic and the PP might be even greater if more than one AP modifies the noun, as in (6b) above.

This data gives support to the claim in Caink (1998, 1999) that Emonds’ Alternative Realization should be revised in order to account for the distribution of pronominal clitics in the Serbian/Croatian/Bosnian (SCB) clause. Consider (25), in which the pronominal clitics *mu* ‘to him’ and *ga* ‘it’ alternatively realise null complement phrases inside VP:

- (25) *Stefan tvrdi da -mu -ga -je Petar_{VP}[poklonio DP[Ø] PP[Ø]]*
 Stefan claims that 3sg.Dat.3sg.Acc. be-3sg. Peter give-ppl.
 ‘Stefan claims that Peter has given it to him as a present’ (Progovac 1996:412)

Setting aside the debate over the exact position of these pronominal clitics, it is clear that they are not in a sisterhood relation with their respective null phrases⁸.

Caink (1998, 1999) therefore argues that Alternative Realization (19) should be reformulated in terms of extended projections rather than sisterhood. UG minimally requires that an AR morpheme may alternatively realize a (possibly null) XP provided the AR morpheme appears within the same extended projection as XP:

- (26) *Revised Alternative Realization: A syntactic feature F matched in UG with category B in the extended projection of Y may be realised in a grammatical morpheme under X⁰, provided X⁰ is in the extended projection of Y⁰.* (Caink 1998:335)

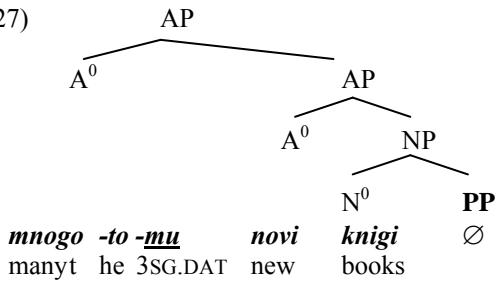
In essence, this is a substantially less restrictive definition of the structural relation that must exist between an AR morpheme and the phrase it alternatively realizes⁹.

Returning to the possessive dative clitic in South Slavic, the tree in (27) illustrates how (26) works in the DP:

⁸ This remains the case even if we take into account Emonds’ more articulated definition of ‘extended sisterhood’ which turns on the fact that purely closed class items are inserted at PF, not in the syntax: ‘if W and Z are sisters, W dominates X, and X dominates the only lexical material under W, then X and Z are sisters’ (Emonds 1999). Space prevents me from discussing ‘extended sisterhood’ in this article, despite the fact that it is a cornerstone of Emonds’ account. However, it is not germane to our concerns in this article.

⁹ See also Lenertová (1999) who argues in favour of (26) on the basis of pronominal clitics in the Czech clause.

(27)



'The many new books'

(The tree in (27) leaves out the DP layer but includes the determiner *-to*). The bound morpheme *mu* ‘his’ appears under A^0 and alternatively realizes the formal features of the null PP. The A^0 is in the extended projection of N^0 (van Riemsdijk 1990, Grimshaw 1991), hence the structural relation between the AR morpheme and the PP is satisfied.

We have already established that an AR morpheme may also ‘double’ the phrase whose formal features it realizes. In fact, Bulgarian avails itself of both options within the DP. In the majority of cases, the possessive clitic licenses a null PP, but consider (28):

- (28) *Majka mu na Ivan*
 mother 3SG.DAT of Ivan
 ‘Ivan’s mother’

In cases of family relations, the clitic doubles the PP in Bulgarian, hence *mu* ‘his’ co-occurs with the PP *na Ivan* ‘of Ivan’.

So far, we have established the mechanism by which the possessive clitic licenses the null PP by introducing a less restrictive Alternative Realization. We now turn to the way in which it is restricted from appearing anywhere in the extended projection of the noun, and thus ensure that the model is not too powerful in its empirical predictions.

4.2 The clitic cluster, Phonological Lexicalization and subcategorization

A significant factor in the theory of Alternative Realization is that semantically null, purely closed class items are inserted at PF (‘Phonological Lexicalization’ Emonds 1985, 1994; Caink 1998). This is motivated by the same argument used by Chomsky (1995) against Agreement Phrases: if an item contributes nothing to LF then it is more economic to exclude it from the syntax. In Emonds’ model, it is cheaper, and therefore obligatory, to insert closed class items at PF (see Caink 1998: chapters 3 and 4 for discussion). Only the presence of ‘purely’ semantic information triggers insertion into the syntax.

It is often argued that pronominal clitics must be inserted into the syntax *en route* for LF because they carry ‘meaning’. However, it should be apparent from the previous sections that pronominal clitics are simply the alternative realization of closed class formal features associated with another node in the tree. The semantics of, say, ‘possession’ is associated with the null PP, and it is the existence of that underlying PP in the syntax that contributes to LF, *not* the presence of the pronominal clitics. This theory therefore makes a substantial break with the traditional view of pronominal clitics as simply unstressed pronouns, and includes them in the typology of lexical items along with verbal inflections, case morphology and null auxiliary verbs such as *be* and *have* in English.

Within an adapted version of Emonds’ theory of lexicalization, Caink (1998) argues that the clitic cluster in the South Slavic clause is formed in the numeration and inserted at PF (according to its subcategorization frame) as a single lexical item. If we employ this same model for the clitic cluster in the DP, then the determiner and possessive clitic are combined in the numeration and inserted together as a single lexical item at PF.

Recall the descriptive generalisation (17), in which we observed that the possessive dative clitic always appears adjacent to the determiner, and otherwise, adjacent to the lexical demonstrative in D^0 . In other words, the possessive clitic always appears adjacent to a morpheme (bound or free) that is specified as [+DEF]. One way of beginning to capture this generalisation is to assume the following subcategorization frame in the lexical entry of the clitic¹⁰:

¹⁰ This was suggested to me by Steven Franks.

(28) *mi*, 1SG.DAT, [+DEF]_____

Clearly, failure to satisfy (28) results in ungrammaticality, as in (13b,c) and (15c,f) above. Moreover, in this theory, the subcategorization frame in (28) drives the formation of the clitic cluster in the numeration: the clitic joins with the AR morpheme marked [+DEF], and then follows the determiner in its placement via strict sisterhood.

Finally, consider again grammatical data that exhibits no determiner, such as (16a), repeated here.

- (30) *Tezi ti knigi*
these 2SG.DAT books
'These books of yours'

The clitic cluster consists of only one member, the possessive dative *ti*. Evidently its subcategorization frame [+DEF]_____ will not have been satisfied in the numeration owing to the absence of an AR morpheme specified as [+DEF]. On insertion at PF, this single member clitic cluster has no other distributional requirement than that of appearing on a head marked [+DEF], hence it is inserted straight onto D^0 , filled by the demonstrative *tezi* 'these'.

5. Summary and Conclusions

This paper has pursued a formal account of the following two descriptive generalisations:

- **Descriptive generalisation #1:** *the determiner is an inflection which appears on the (head of the complement to D^0)*
- **Descriptive Generalisation #2:** *The possessive dative clitic appears adjacent to any determiner; if none is present, it appears adjacent to the lexical demonstrative in D^0 .*

I argued that both the South Slavic bound determiners/demonstratives and the possessive dative clitic are PF-inserted morphemes that alternatively realize the formal features on nodes elsewhere in the tree. We have seen that the bound determiner/demonstrative morpheme is the realization of the [+DEF] feature in D^0 . It is alternatively realized as predicted by Emonds' Alternative Realization via sisterhood onto the head of the complement to D^0 . This complement may be the NP or the first of a number of modifying APs. The possessive clitic is the alternative realization of the formal features ([DATIVE] and Φ-features) associated with a null possessor PP in N'.

In both cases, the alternative realization morpheme may 'double' an overt node. Hence, (i) in colloquial Bulgarian, the determiner may co-occur with demonstratives and (ii) in DPs expressing family relations, the possessive clitic obligatorily co-occurs with a possessive postnominal PP. In this theory, such doubling is formally equivalent to 'clitic doubling' in the Macedonian clause.

Both morphemes are subject to Phonological Lexicalization (or 'PF-insertion') because neither morpheme carries any semantics; rather, each alternatively realizes a (possibly null) node elsewhere in the tree that is itself the bearer of semantic information for LF.

Having established the precise nature of these lexical items in a way that is rarely done in the generative literature, I then addressed their inter-related distribution. I argued that the strict adjacency between the determiner and the clitic follows straightforwardly from the fact that they are members of a clitic cluster, formed in the numeration and inserted as a single lexical item at PF. The possessive dative clitic has the subcategorization frame [+DEF]_____ which is satisfied in the numeration by combining with the determiner morpheme. If no determiner is present, the possessive clitic is lexicalized alone, again according to the subcategorization requirements. In this way, it may appear adjacent to a demonstrative. If neither a determiner nor demonstrative is present, the subcategorization frame remains unsatisfied and the DP is ungrammatical.

Observing that the possessive dative clitic is not necessarily in a sisterhood relation with the null PP in N', we adopted Caink's (1998) revised Alternative Realization whereby a morpheme may license a null node provided it appears in the same extended projection as that node.

In this way, the paper presents an independently motivated theory of closed class items to account for the nature and inter-related distribution of the South Slavic determiner and possessive clitic.

References

- Abney, S. (1987) 'The English Noun Phrase in its Sentential Aspect', Ph.D. Diss., MIT, Cambridge, Mass.
- Arnaudova, O. (1996) 'The Bulgarian Noun (Adjective) Movement to D', *Trondheim Working Papers in Linguistics* vol. 28:1-28.
- _____(1998) 'Demonstratives and the structure of the Bulgarian DP', *Trondheim Working Papers in Linguistics* 31:1-12.
- Caink, A. (1998) 'The lexical interface: closed class items in South Slavic and English', Ph.D. diss., University of Durham.
- _____(1999) 'The Serbian/Croatian/Bosnian clitic cluster at the lexical interface', in *Formal Approaches to Slavic Linguistics* vol. 7, edited by Dziwirek, K., H. Coats, & C. Vakareliyska. Ann Arbor: Michigan Slavic Publications: 81-100.
- Dimitrova-Vulchanova, M. & G. Giusti (1999) 'Possessors in the Bulgarian DP', *Topics in South Slavic Syntax and Semantics*. Amsterdam: John Benjamins.
- Emonds, J. (1985) *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.
- _____(1987) 'The Invisible Category Principle', *Linguistic Inquiry* 18:613-632.
- _____(1994) 'Two Principles of Economy' in Cinque, Koster, Pollock, Rizzi, & Zanuttini (eds.) *Paths towards Universal Grammar*. Georgetown University Press: 85-110.
- _____(1999) 'How clitics license null phrases; a theory of the lexical interface', in Riemsdijk, H.C van (ed.). *Empirical Approaches to Language Typology: clitics in the languages of Europe*. The Hague: Mouton de Gruyter.
- Ewen, R. (1979) 'A grammar of Bulgarian clitics', PhD diss., University of Washington.
- Fowler, G. & S. Franks (1994) 'DP in Bulgarian', in *Indiana Slavic Studies* 7:58-68.
- Franks, S. (1998) 'Clitics in Slavic', Position paper, Comparative Slavic Morphosyntax workshop, Spencer, Indiana (<http://www.indiana.edu/~slavconf/linguistics/index.html>).
- Grimshaw, J. (1991) 'Extended Projections', ms. Brandeis.
- Halpern, A. L. (1995) *Topics in the Placement and Morphology of Clitics*. Stanford: CSLI publications.
- Lenertová, D. (1999) 'Internal Topics and CP-structure in Czech', paper presented at *Formal Description of Slavic Languages III*, University of Leipzig.
- Lunt, H. (1952) *A Grammar of the Macedonian Literary Language*. Skopje: Dr. Knigoizdatelstvo na NR Makedonija.
- Nikanne, U. (1993) 'The semantic case endings in Finnish', in Holmberg, A. & U. Nikanne (eds.) *Case and Other Functional Categories in Finnish Syntax*. The Hague: Mouton de Gruyter.
- Progovac, L. (1996) 'Clitics in Serbian/Croatian: Comp as the Second position', in Halpern & Zwicky (eds.) *Approaching Second Position: second position clitics and related phenomena*. CSLI Publications: 411-428.
- Riemsdijk, H. van (1990) 'Functional Prepositions' in Pinkster, H. & I. Genee (eds.) *Unity in Diversity: papers presented to Simon C. Dik on his 50th birthday*. Dordrecht: Foris.
- Rudin, C. (1997) 'AgrO and Bulgarian pronominal clitics', in Lindseth, M. & S. Franks (eds.) *Formal Approaches to Slavic Linguistics* vol. 5. Ann Arbor: Michigan Slavic Publications: 224-252.
- Scatton, E. (1984) *A Reference Grammar of Modern Bulgarian*. Columbus, Ohio: Slavica.
- Zec, D. & S. Inkelaas (1990) 'Prosodically constrained syntax' in Zec, D. & S. Inkelaas (eds.) *The Phonology-Syntax Connection*. University of Chicago Press.