
A Domain-Based Approach to 2P Clitics in Pashto

ASCANDER DOST

1. Overview

Second position (2P) clitic phenomena have provided contemporary linguistics with a rich testing ground for the autonomy of grammatical components, as their placement and ordering often seems to require reference to both syntax and prosody. 2P clitics in Pashto frame these questions in a particularly interesting way; the constraints governing their placement cut across both syntax and prosody, and *at both* the lexical and sublexical levels. In this paper, I discuss a recent analysis of Pashto 2P clitics in Roberts (2000) and address some issues facing the proposal. I then present an alternative analysis within the formalism of Head-Driven Phrase Structure Grammar (Pollard and Sag, 1987, 1994, Sag and Wasow, 1999, Ginzburg and Sag, 2001, henceforth, HPSG), augmented with the word order domains of Kathol (1995).

2. Background

Pashto 2P clitics include pronominal, verbal and adverbial elements (see FIGURE 1) and display properties commonly associated with postlexical clitics; namely they are prosodically dependent on an adjacent prosodic element, exhibit a low degree of selection with respect to the syntactic category of their hosts, and take wide scope with respect to coordinated hosts (see Zwicky and Pullum (1983) for further discussion of the clitic/affix distinction). Tegey (1977) formulates the generalisation that 2P clitics appear after the first *stress-bearing* phrasal constituent

<i>Pronominal</i>		<i>Modal</i>	<i>Adverbial</i>
me	1SG	ba FUT, <i>can, may</i>	xo <i>indeed, really</i>
de	2SG	de <i>should, must</i>	no <i>then</i>
ye	3SG, 3PL		
am, mo	1PL, 2PL		

FIGURE 1 Inventory of Pashto 2P clitics

in the Pashto clause. As the examples in (1) illustrate, placement of 2P clitics may not result in discontinuous syntactic phrases (clitics here and throughout shown in boldface).

- (1) a. [aʃa xɑysta p y la]_{NP} **me** w lida
that pretty girl cl.1SG PF.see.PST.3SG.FEM
I saw that pretty girl.
- b. *aʃa **me** xɑysta p y la w lida
- c. [p  ʈop k]_{PP} **ye** zmaray wuw ft 
with gun cl.3SG lion PF.shoot.PST.3SG
He shot a lion with a gun.
- d. *p  **ye** ʈop k zmaray wuw ft 

Furthermore, the phrasal host must be stress-bearing; it must contain at least one primary accent. Unaccented constituents may not host 2P clitics, as the example in (2) illustrates.

- (2) a. [rata]_{PP} [pe]_{PP} [gand ]_V **de**
for-me by-him sew.PF.PST.3SG cl.2SG
You were having him sew (it) for me. (Hock 1996, p. 235)
- b. *[rata]_{PP} **de** [pe]_{PP} [gand ]_V

3. A Derivational Approach

A recent analysis of Pashto 2P clitics in Roberts (2000), articulated within the framework of the Minimalist Program (Chomsky, 1995), takes the position that (pronominal) 2P clitics are agreement morphemes. This assumption is based on the observation (also made in Babrakzai (1999)) that pronominal 2P clitics are in complementary distribution with verbal agreement morphology. That is, only ergative and accusative arguments may be cliticised. Nominative or absolutive arguments cannot. Roberts claims (following Sportiche (1996)) each clitic heads an agreement projection, whose specifier licenses a null pronominal argument. Under the analysis, 2P clitics are base-generated in high clausal position. The second position effect is derived through obliga-

tory syntactic movement. For example, the sentence in (3) receives the phrase structure analysis in FIGURE 2.

- (3) [spín]_{NP} **me** wənəlídə
 spin cl.1SG PF.NEG.SEE.PST.3SG.MASC
I didn't see Spin. (Roberts 2000, p. 81)

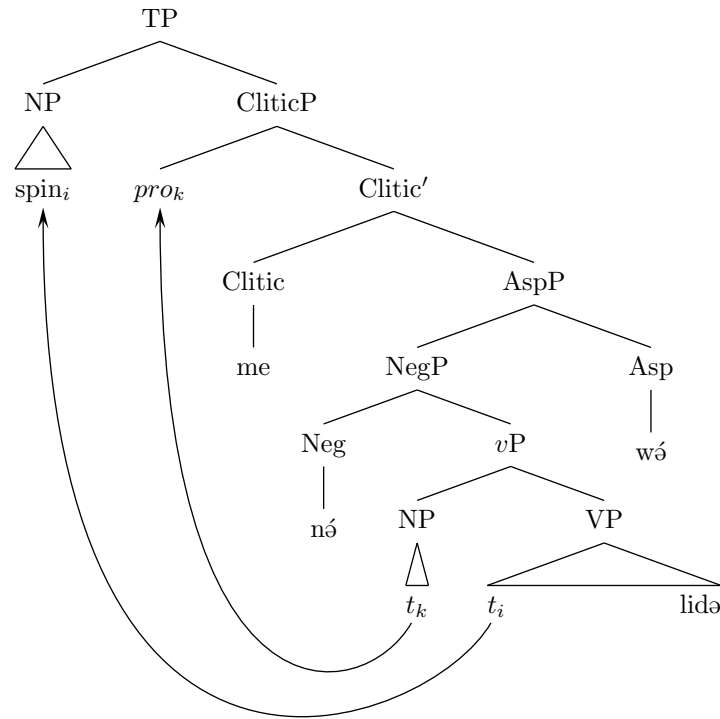


FIGURE 2 Phrase structure analysis of (3)

In the structure above, the pronominal clitic **me** is base-generated as the head of CliticP. The second position effect is derived by movement of the NP **Spin**, triggered by the EPP requirement that [Spec TP] be filled at Spell-Out (Roberts 2000, p. 83).

In Pro-drop clauses that contain only a verb and 2P clitics (henceforth, VC clauses), however, the EPP fails to trigger movement of the verb. As noted in Roberts (2000), a purely syntactic account incorrectly derives ungrammatical orderings. As the examples in (4) indicate, 2P clitics in VC clauses are hosted by the verb.

- (4) a. pórebdə **me**
 beat.IMPF.PST.3SG cl.1SG
 I was beating him. (Tegey 1977, p. 88)
- b. ***me** pórebdə
- c. likóm **ye**
 write.IMPF.PR.1SG cl.3SG
 I am writing it. (Babrakzai 1999, p. 93)
- d. ***ye** likóm

Roberts argues for a “last resort” application of Prosodic Inversion (Halpern, 1995, henceforth PI), which inverts 2P clitics with prosodic words on their right, to repair faulty syntactic derivations, as in FIGURE 3. At the post-syntactic level of PF, PI applies to repair syntactic

$$[\text{ye}] [\omega \text{ likóm}] \Rightarrow [\omega [\omega \text{ likóm}] \text{ ye}]$$

FIGURE 3 Prosodic Inversion in (4c)

structures in which the 2P clitic is left without an appropriate prosodic host.

To summarise briefly, the approach proposed in Roberts (2000) argues that (pronominal) 2P clitics in Pashto are agreement morphemes base-generated in high clausal position. The second position placement of clitics within the clause is derived primarily in the syntax, with obligatory movement of an overt phrase to [Spec, TP], to satisfy the EPP. In VC clauses, the EPP does not force movement, and the offending structure is repaired post-syntactically by an application of PI which inverts 2P clitics with the nearest prosodic word.

3.1 Issues facing the proposal

The proposal in Roberts (2000) faces three main issues that I discuss in this section. First, the *clitics-as-agreement* hypothesis makes the prediction that 2P clitics should double overt pronominal arguments, which they do not. Secondly, the PI strategy that derives examples like (4a,c) encounters difficulty with several cases of apparent clitic infixation. Third, the intuition that 2P clitics are *prosodically* required to have hosts is lost by the primarily syntactic nature of the proposal. That is, the trigger for movement in general cases (the EPP) is a syntactic requirement and not a prosodic one.

The hypothesis that 2P clitics in Pashto are actually agreement morphemes suggests that they should cooccur with full nominal arguments. This prediction is not borne out in the data, as shown by (5).

- (5) a. zə kɪtəb əxləm
 I book buy.IMPF.PR.1SG
I'm buying a book.
- b. *zə **me** kɪtəb wəxlə
 I cl.1SG book PF.buy.PST.3SG.MASC
I bought a book.

In (5a) the first person pronoun **zə** is doubled by first person inflection on the verb **əxləm**. The expectation is that the 2P clitic **me** should similarly double the pronoun, which (5b) shows not to be the case. Roberts (2000) addresses the issue and suggests that an explanation of the absence of 2P clitics doubling full NPs may be available if one assumes (following Jaeggli (1982)) that clitics absorb Case, rendering overt NPs without a means to satisfy the Case Filter.

A second issue facing the analysis in Roberts (2000) concerns several cases of apparent clitic infixation. Tegey (1977) identifies three classes of main verbs in Pashto that sometimes require 2P clitics to appear after the first accented syllable of the verb, rather than after the verb itself. Example (6) illustrates one such case, conditioned on the aspect of the verb **ʈəlwəhə**.

- (6) a. ʈəlwəhə́ **me**
 push.IMPF.PST.3SG cl.1SG
I was pushing it. (Tegey 1977, p. 92)
- b. *ʈəl **me** wəhə́
- c. ʈəl **me** wəhə
 ? cl.1SG push.PF.PST.3SG
I pushed it. (Tegey 1977, p. 92)
- d. *ʈəlwəhə **me**

Examples (6c,d) show that when stress occurs on the first syllable of the verb (in the perfective aspect), 2P clitics obligatorily occur after the stressed syllable. The issue facing a PI account of clitic placement is how to appropriately define the prosodic element that inverts with the clitic. Cases like (6) suggest that a finer-grained notion of prosodic constituency than is traditionally assumed must be used to account for the Pashto facts.

A third issue concerns the conceptual awkwardness of the “syntax first—then prosody” derivation of the second position effect. The analysis seeks to derive second position by means of a purely syntactic movement (EPP). It is only if EPP movement fails to apply that prosodic

considerations are brought to bear on clitic placement. This seems to conflict with the intuition that clitics are prosodically dependent and “lean” on some adjoining host element. Note that the *clitics-as-agreement* approach does away with clitics as a natural class. However, the issue concerns 2P clitics empirically, regardless of what theoretical status they have. The division of labour, so to speak, for ensuring clitics have hosts does not match well with general observations about their prosodic dependencies.

To summarise briefly, the approach taken in Roberts (2000) faces three issues worth pointing out. First, identifying clitics with agreement morphemes makes an incorrect empirical prediction about their behaviour with respect to clitic doubling. Secondly, it is unclear how the PI strategy is to be constrained such that examples like (4a) and (6c) are uniformly predicted. Finally, the split nature of how clitic placement is derived clashes with the intuition that clitics have prosodic demands of their hosts. The data suggest a unified treatment of clitic positioning would be more appropriate, which I explore in the following sections.

4. A Domain-Based Approach

4.1 Architecture

HPSG is a constraint-based theory of grammatical competence that relies on two fundamental components: a highly structured representation of grammatical categories, encoded as typed feature structures, and a set of descriptive constraints on types, which restricts the expressions admitted as part of a given natural language (Levine and Meurers, to appear). The set of descriptive constraints are declarative and unordered, making HPSG non-derivational. Formally, an HPSG grammar is given by a set of *type constraints* and a *signature*. The signature consists minimally of an enumeration of the set of grammatical *types* (also called *sorts*), as well as a statement of which features are appropriate for each type and a statement of what type of value is appropriate for each feature (Ginzburg and Sag 2001, p. 17). The basic unit of currency in HPSG is the *sign*, which encodes linguistic information such as the syntactic/semantic, phonological and contextual properties of grammatical objects. Signs are represented as Attribute-Value Matrices (AVMs) of the sort in FIGURE 4.¹ Both word-level and phrase-level objects are analysed in terms of *signs*. Syntactic/semantic information

¹I will employ the following notational conveniences throughout this paper: crucial aspects of AVMs will be highlighted **as so**. Coindexed boxed numerals (like $\boxed{1}$) indicate structure-sharing of feature content. For space reasons, features and values will be abbreviated where possible.

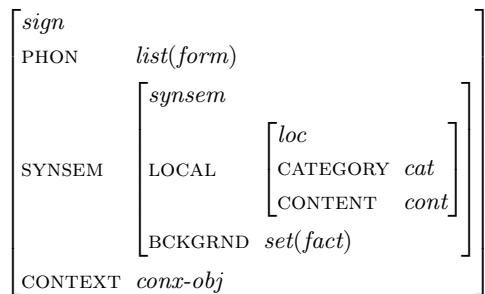


FIGURE 4 (Part of) the structure of *sign*

is encoded on the feature SYNSEM, and phonological/prosodic properties are encoded on the feature PHON(OLGY). The type *synsem* may be viewed as an analogue to traditional syntactic category representations (e.g., N, V, Adj). Phrasal *signs* are licensed by type constraints on well-formed objects. For example, the Generalised Head Feature Principle (GHFP, see (7) below) licenses instances of *h(ea)d(ed)-ph(rase)* such that the SYNSEM value of the mother is structure-shared (i.e., token-identical) with the SYNSEM value of exactly one head daughter. This is analogous to the ‘X’ identity condition of \bar{X} Theory (Ginzburg and Sag 2001, p. 33).

- (7) a. *hd-ph*:

$$\left[\text{SYNSEM} \quad / \square \right] \rightarrow \dots \mathbf{H} \left[\text{SYNSEM} \quad / \square \right] \dots$$

a'. Headed phrases contain exactly one head daughter.

Additional type constraints that will be relevant to the present discussion are given explicitly in (8) below.

- (8) a. *hd-comp-ph*:

$$\left[\right] \rightarrow \mathbf{H} \left[\begin{array}{l} \textit{word} \\ \text{COMPS} \quad \textit{nelist}(\square \oplus \textit{list}) \end{array} \right], \square$$

a'. Head-complement phrases (e.g., VP) contain a lexical head and its complements.

- b. *hd-subj-ph*:

$$\left[\text{SUBJ} \quad \langle \rangle \right] \rightarrow \left[\text{SYNSEM} \quad \square \right], \mathbf{H} \left[\begin{array}{l} \text{SPR} \quad \langle \rangle \\ \text{SUBJ} \quad \langle \square \rangle \end{array} \right]$$

b'. Head-subject phrases (e.g., S) contain a head daughter and exactly one subject.

To illustrate with a small example, the English sentence **Randi likes kittens** receives the phrase structure analysis in FIGURE 5. The verb

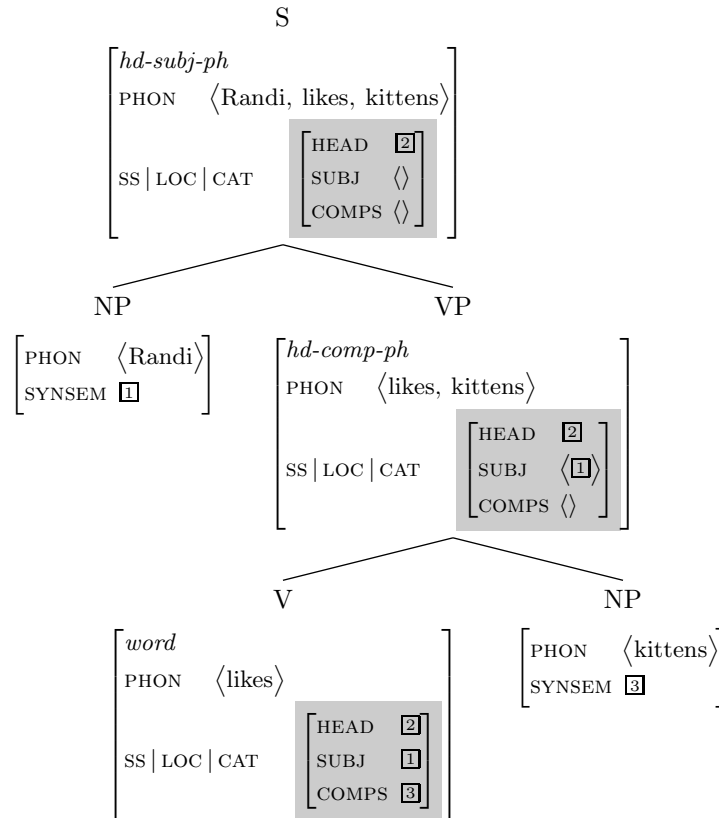


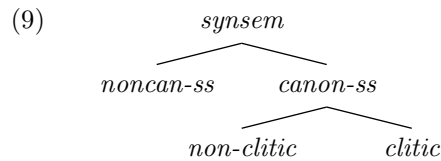
FIGURE 5 (Partial) syntactic analysis of **Randi likes kittens**

likes subcategorises for two arguments: the object NP **kittens** (by way of the feature COMPS) and the subject **Randi** (by way of SUBJ). The VP (an instance of *hd-comp-ph*) structure shares its HEAD features with those of its head daughter (V). Note that at the VP level, the COMPS value of the main verb has been saturated. The clause (an instance of *hd-subj-ph*) is (partially) licensed by the type constraint in (8b). The VP combines with the subject **Randi**. Again, the SUBJ feature on the mother node bears an empty list, corresponding with argument saturation. A grammatical analysis of the sentence emerges through the interaction of lexical entries for the terminal elements, and type

constraints that mediate syntactic structure.

4.2 The syntax of 2P clitics

The intuition I wish to pursue in this section (originally proposed in Kupść (2000) for Polish clitics) is that the clitic *vs.* non-clitic distinction is orthogonal to the syntactic properties of lexical *signs*. That is, the distinction between clitics and non-clitics is one that runs parallel to the purely syntactic aspects of lexical items. Hence, an (*e.g.*, pronominal) clitic is the same syntactic category as its plain-word analogue, but represents a different class of *synsem* object. Following Kupść (2000), I adopt the partition of the type *canonical-synsem* into *clitic* and *non-clitic* subtypes, as shown in (9).



Setting aside for the moment the featural differences between *clitic* and *non-clitic* types², the approach I take here treats clitics identically with their non-clitic counterparts from a syntactic point of view. The same constraints that license overt pronominal arguments in Pashto clauses will also license pronominal 2P clitics, and similarly for modal and adverbial clitics. To illustrate, the clitic **me** in (3) has the lexical entry in FIGURE 6 ((3) repeated here as (10)).

- (10) spín **me** wənəlīdə
 spin cl.1SG PF.NEG.see.PST.3SG.MASC
I didn't see Spin.

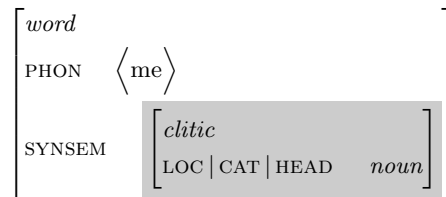


FIGURE 6 (Partial) lexical entry for **me**

²Simply, the distinction between *clitic* and *non-clitic* SYNSEM types is based on prosodic deficiency. Clitics subcategorise for a prosodic host, while non-clitics do not. Interested parties should consult Klein (2000) for a constraint-based approach to prosody.

The lexical entry in FIGURE 6 straightforwardly captures the lack of clitic doubling in Pashto. As a nominal *sign*, the clitic **me** occurs as an argument to a main verb. That is, the analysis here claims that pronominal clitics are arguments and not agreement morphology, so there is no expectation that the clitic should double overt NPs. The example in (10) is licensed by existing type constraints on phrasal *signs*, and receives the phrase structure analysis in FIGURE 7. Nothing additional

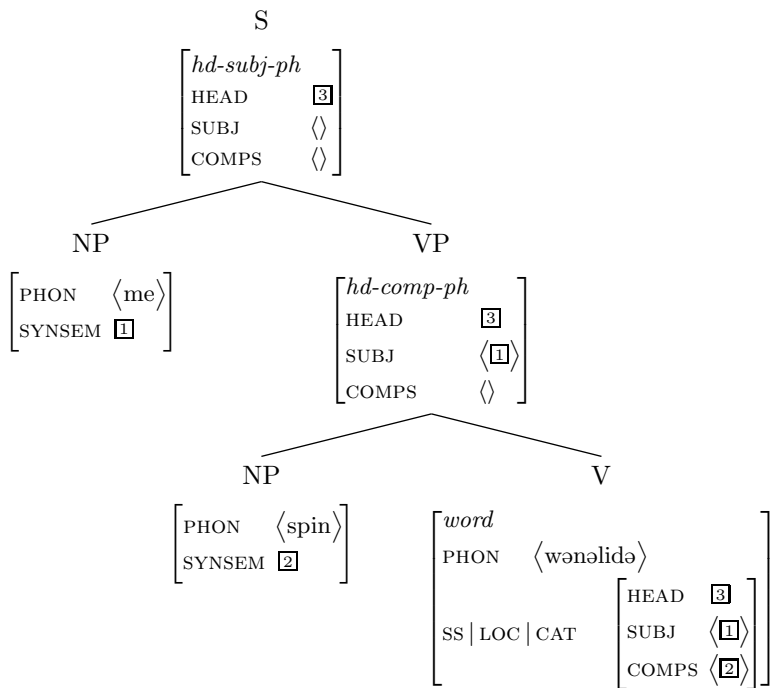


FIGURE 7 (Partial) syntactic analysis of (10)

needs to be said about how clauses containing pronominal 2P clitics are syntactically licensed. It is also important to mention that the syntactic analysis laid out here makes no claims about word order in the Pashto clause. The phrase structure in FIGURE 7 is an analysis of the hierarchical structure of example (10). The linear ordering of elements in a clause is mediated by the domain ordering component of the grammar. In the next section, I develop an account of ordering effects in Pashto clauses.

4.3 Word order domains

Domain-based approaches to linear ordering phenomena within the HPSG literature (Reape, 1994, Kathol, 1995, Penn, 1999, among others) represent one of several strands of research that depart from traditional notions regarding the relationship between the syntactic structure and linear organisation of strings (see Kathol (1995) for further discussion of the history of domain-based approaches to word order phenomena). The intuition behind word order domains is that constraints on hierarchical structure are related to, but formally distinct from those on linear realisation, a claim made as early as Curry’s (1961) distinction between *tectogrammatical* structure, where meanings are constructed compositionally, and *phenogrammatical* structure, where constraints on the surface realisation of a string are resolved (see also Dowty (1982, 1996)).

In HPSG approaches to word order domains, the ordering of elements that make up the terminal yield of a *sign* (word order) is encoded on a *sign* level attribute $\text{DOM}(\text{AIN})$. A general constraint requires the phonological yield of a sign to be equivalent to the concatenation of the PHON values of objects on its DOM list, as shown in FIGURE 8. The



FIGURE 8 Constraint on the phonological yield of *sign* (Reape, 1993)

constraint in FIGURE 8 simply requires the PHON value of a sign to correspond with the PHON values of each of its domain elements.

The analysis I propose here also makes use of the notion of *topological fields*, as used in Kathol (1995). Briefly, topological fields are position classes that model the linear distributional properties of syntactic elements, without making claims about their hierarchical organisation.³ The intuition is that word order patterns of a language can be described in terms of membership in (linearly) ordered fields. Kathol (1995) makes formal use of such fields to account for word order effects in German clauses.

The intuition behind the analysis of 2P clitic placement in Pashto clauses is that the linear ordering of a grammatical object is licensed in

³Interested parties should consult Höhle (1986) for a review of topological fields theory.

tandem with its hierarchical structure. Constraints on word order and those on hierarchical organisation operate in parallel. Additionally, the surface constraints on linear order are argued to result from interactions between topological fields, syntactic constraints, and prosodic requirements.

As a starting point, I assume the topological model for simple Pashto clauses in FIGURE 9. In the proposed topology, elements that host 2P

	Pre-clitic (<i>pre-cf</i>)	Clitic (<i>cf</i>)	Middle (<i>mf</i>)	Verbal (<i>vf</i>)
a.			zə kɾɪɒb I book	axləm buy.IMPF.PR.1SG
b.	spín spín	me cl.1SG		wónəlídə PF.NEG.see.PST.3SG
c.	likám write.IMPF.PR.1SG	ye cl.3SG		

FIGURE 9 Topological spaces in simple Pashto clauses

clitics instantiate the Pre-clitic (*pre-cf*) field, while 2P clitics occur in the Clitic (*cf*) field. The Middle (*mf*) field is the default field assignment for the clause, and verbal elements instantiate the Verbal (*vf*) field, anchoring the edge of the clause. This division of fields is motivated in part by grammatical variations between OSV and SOV word order, illustrated in (11).⁴

- (11) a. spi pɪʃo xog kɾa
dog cat hurt do.PST.PF.3SG.FEM
The dog hurt the cat. (Roberts 2000, p. 12)
- b. pɪʃo spi xog kɾa
cat dog hurt do.PST.PF.3SG.FEM
The dog hurt the cat.

Following Kathol (1995), I formalise the notion of topological fields as a type *topo(logical)-field*, with the hierarchy in FIGURE 10. A few important things to note about the hierarchy of topological field types; HPSG feature descriptions are maximal, therefore all field assignments must be resolved. The nonmaximal types I assume are *verbal* and *nonverbal*. The *verbal* type dominates both *pre-cf* and *vf* assignments. Intuitively,

⁴OSV word order in Pashto clauses is more complicated than I wish to address in this paper. Specifically, OSV ordering is permissible when inflectional morphology on the verb disambiguates grammatical role. In the absence of suitable inflection (i.e., in present tense clauses), OSV ordering requires a heavy pause after the direct object.

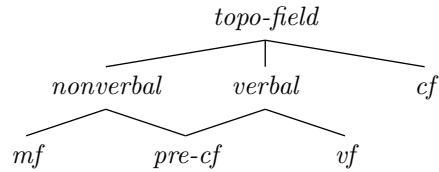


FIGURE 10 Hierarchy of topological field types

verbal lexical items can either occur clause finally (as in (1)) or can host clitics (as in (4)). Similarly, *nonverbal* field assignments can resolve to either *mf* (general case) or *pre-cf* (when hosting 2P clitics).

As mentioned earlier, the constraints on 2P clitic placement arise from the interaction of relatively simple constraints in different areas of the grammar. I make use of two constraints on topological fields, given in (12).

- (12) a. **Linear Precedence:**
 $pre-cf \prec cf \prec mf \prec vf$
 a'. Topological fields occur in a fixed order.
- b. **Field Uniqueness:**
 $pre-cf \prec pre-cf$
 b'. The *pre-cf* field must be unique (*i.e.*, contain a single member).

Linear Precedence requires typed topological fields to occur in a specific order. Field Uniqueness constraint ensures that in any Pashto clause, only one object may instantiate the *pre-cf* field.

Again, following Kathol (1995) I assume that syntactic constraints on field assignments may be either lexical or constructional. That is, lexical constraints condition field assignments based on lexical class, while constructional constraints condition field assignments in phrasal *signs*. I propose that *cf* and *verbal* field assignments are lexical, applying to 2P clitics and verbs, respectively. The relevant constraints are given in (13).

- (13) a. $clitic \Rightarrow \left[\text{DOM} \left\langle \left[\begin{array}{c} cf \\ PHON \quad \dots \end{array} \right] \right\rangle \right]$
- a'. 2P clitics are typed *cf*.
- b. $verb \Rightarrow \left[\text{DOM} \left\langle \left[\begin{array}{c} verbal \\ PHON \quad \dots \end{array} \right] \right\rangle \right]$
- b'. Verbs are typed *verbal*.

Since the domain of the present analysis is restricted to simple Pashto clauses, a constraint on field assignment in headed phrases (*hd-ph*) is given in (14).

- (14) **Domain Condition:** (Informal version)
 In all instances of headed phrase (*hd-ph*) such that:
- (a) the head daughter is typed *verbal*
 - (b) all non-head daughters contain *non-clitic* SYNSEM values
- the head daughter is typed *vf* and all non-head daughters are typed *nonverbal*.

The constraint above targets phrases headed by a *verbal* element and containing no 2P clitics. Such phrases are constrained to resolve the head to the *vf* field assignment and all non-head daughters to the *non-verbal* assignment. The constraint has the effect of ensuring that verbs occur on the right edge of the clause, in all cases where there is overt phonological material to host 2P clitics.

Returning to the distinction between *clitic* and *non-clitic* synslems, I propose a prosodic constraint on all 2P clitics requiring them to have phonological hosts. Clitics are distinguished from non-clitics with respect to prosodic dependency. Intuitively, instantiation of the *pre-cf* field can be seen as a reflex of this requirement. That is, the second position effect under the current analysis receives a unified explanation in terms of prosodic subcategorisation.

- (15) **Clitic Condition:** (Informal version)
 2P clitics subcategorise *full* prosodic structures to their immediate left.⁵

⁵Here, *full* denotes a type of prosodic structure in Klein (2000) that subsumes both prosodic words and larger metrical structures. In the interest of space, I refrain from a full discussion of the analysis of prosodic structure in Pashto. The crucial point I wish to make here is only that prosodic dependencies can be modeled as well formedness conditions on structures that contain clitics. The current analysis is compatible with a number of ways of modeling prosodic dependency in HPSG.

To illustrate how the analysis here accounts for the various 2P clitic placement patterns in Pashto, I present three representative examples and discuss how the analysis predicts each. First, consider the now-familiar example in (3) and (10) (in its third incarnation, as (16)). Under the present approach, the sentence receives the domain ordering analysis in FIGURE 11.

- (16) *spín me wənəlīdə*
 spin cl.1SG PF.NEG.see.PST.3SG.MASC
I didn't see Spin.

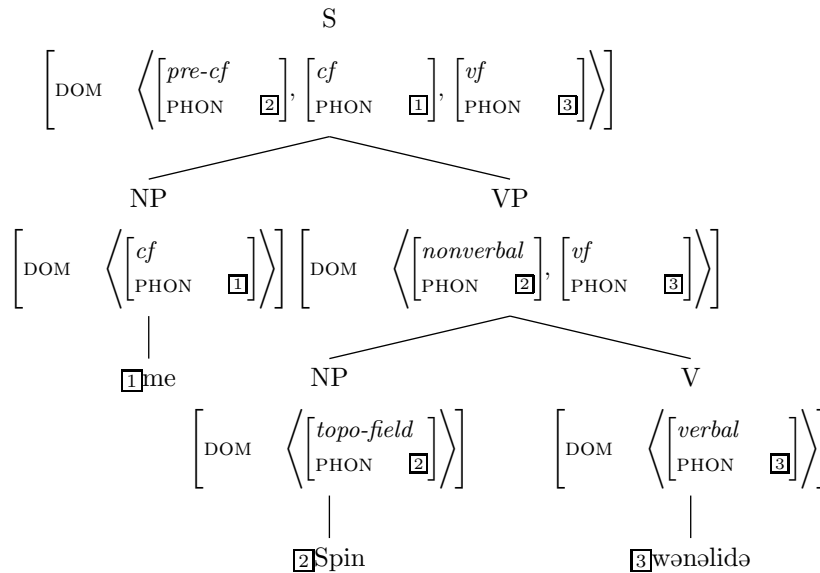


FIGURE 11 Ordering analysis of (16)

The lexical constraints on field assignment in (13) ensure that the verb *wənəlīdə* is typed *verbal* and the 2P clitic *me* is typed *cf*. This means the verb must resolve to either *vf* or *pre-cf* at the level of the clause (see FIGURE 10). At the level of the VP, the Domain Condition applies, and the verb is resolved to *vf*, while the NP *Spin* resolves to *nonverbal*. Finally, at the S-level, the Clitic Condition forces a resolution of the object NP to *pre-cf*, satisfying the prosodic dependency of the clitic. Field ordering is enforced throughout by Linear Precedence and Field Uniqueness.⁶

⁶While it seems intuitive to think of field resolution in terms of derivations, there

Consider now examples such as (4), which are accounted for in Roberts (2000) by means of a post-syntactic PI movement. The example in (4c) (repeated here as (17)) receives the domain ordering analysis in FIGURE 12.

- (17) **likəm** **ye**
 write.IMPF.PR.3SG cl.3SG
I am writing it.

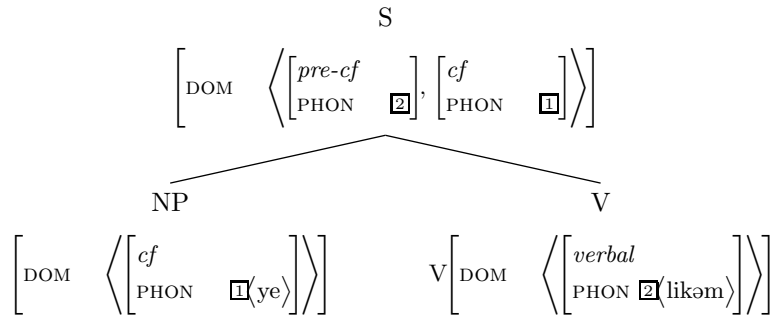


FIGURE 12 Ordering analysis of (17)

As in the previous example, the verb **likəm** and 2P clitic **ye** are typed *verbal* and *cf* respectively, by virtue of the lexical constraints in (13). The resulting phrase is not subject to the Domain Condition, as it contains a *clitic* typed SYNSEM value. At the topmost level, the Clitic Condition compels a resolution of the *verbal* domain object to *pre-cf*, to ensure the clitic has a host. As before, the ordering of fields and uniqueness of the *pre-cf* field are mediated by the constraints in (12). In the analysis proposed here, the difference between cases like (1) and (4) has to do with application of the Domain Condition. That the 2P clitic has a host in both cases receives a uniform explanation in terms of prosodic dependency.

Finally, I address examples like (6), in which the 2P clitic appears to split lexical items. In order to account for these data, I make the additional assumption (following Crysmann (2000)) that word-level entities may contribute more than one domain object.⁷ That is, verbs like

is nothing derivational about the current account. Resolutions of topological fields to any other assignments render the string ungrammatical. That is, there is exactly one assignment that corresponds with a grammatical clause. For expository clarity, explanations of domain analyses may have this flavour.

⁷The analysis in Crysmann (2000) targets subject agreement markers in Udi, which display nontrivial similarities to 2P clitic placement in Pashto. Specifically,

təlwahə in Pashto are syntactic atoms, but are complex from an ordering perspective. To illustrate, the main verb in (18) (repeated from above) has the lexical entry in FIGURE 13.

- (18) **təl me wahə**
 ? cl.1SG push.PF.PST.3SG
I pushed it.

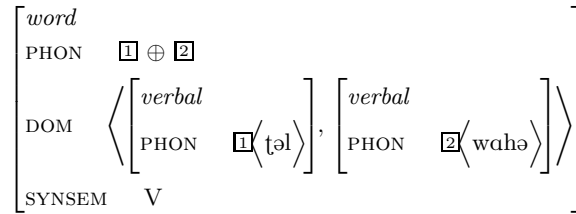


FIGURE 13 (Partial) lexical entry for **təlwahə**

The lexical entry in FIGURE 13 captures the fact that while the verb itself constitutes a syntactic atom (its SYNSEM value is analogous to all main verbs), it remains separable from a domain ordering perspective. That is, the lexical item contains more than one domain object on its DOM list. The example in (18) receives the domain ordering analysis in FIGURE 14. Here, both domain objects associated with the verb are

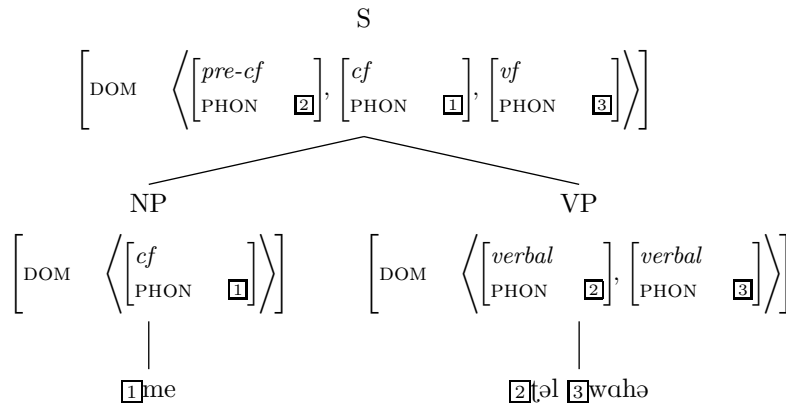


FIGURE 14 Ordering analysis of (18)

typed *verbal*. As in the previous example, the Domain Condition does agreement markers may be realised verb-internally under certain circumstances.

not apply (the phrase contains a *non-clitic* valued SYNSEM feature). The Clitic Condition forces a resolution of the domain object associated with **tə**l to the *pre-cf* field. Field Uniqueness ensures that only one *pre-cf* typed element occurs in the clause. The remaining element must resolve to *vf* in order for the clause to be licensed.

To summarise, in the analysis of their syntactic licensing, 2P clitics are treated as separate classes of *synsem*, capturing the intuition that they are different from their non-clitic analogues in some respects (prosodic deficiency), and not in others (occur as arguments to verbs, etc.). The analysis requires nothing new to be said about how syntactic objects containing 2P clitics are licensed. Furthermore, since pronominal clitics are treated as arguments, the analysis avoids the prediction that they should double overt NPs. The analysis of Pashto clause ordering assumes a topology for simple clauses in which the verb anchors the right edge of the clause. Clitic hosts instantiate the *pre-cf* field, clitics occur in the *cf* field, and the default field assignment is *mf*. The intuition pursued by the analysis is that word order effects are constrained in parallel with syntactic structure. Word order is argued to result from the interaction of multiple grammatical components, including syntax and prosody. Assuming the topological model of simple Pashto clauses in FIGURE 9, and the hierarchy of topological types in FIGURE 10, the proposed analysis accounts for 2P clitic placement through the interaction of simple constraints. Topological constraints mediate field ordering and uniqueness. Lexical constraints specify field assignments that are present in the lexical entries of both 2P clitics and verbs. The Domain Condition mediates field assignment in headed phrases, and the Clitic Condition is a prosodic requirement on 2P clitics that is responsible for “deriving” the second position effect. The immediate benefits of the current proposal include a prediction of the absence of clitic doubling, a unified treatment of the second position effect in terms of prosodic subcategorisation, and an account of apparent clitic infixation that avoids the complications of a PI-based account.⁸

5. Remarks and Conclusions

Some issues remain for the analysis presented in this paper that merit brief discussion here. First, the analysis does not explicitly address the ban on pronominal 2P clitics from appearing as objects of adpositions. Moreover, this seems to be a general ban cross-linguistically. A solution

⁸The Clitic Condition merely requires the host of a 2P clitic to be of the prosodic type *full*. This underspecification allows for the clitic’s host to range from larger metrical structures to prosodic words, and in the case of examples like (6), sublexical prosodic units.

along the lines of prosodic incompatibility is suggested by the proposal here. Assuming the Clitic Condition requires 2P clitics to have suitable prosodic hosts, the ban on such clitics appearing in PPs receives an explanation if one argues that adpositional elements in Pashto are ill-suited for this role. That is, elements of category P do not constitute *full* prosodic objects. A second issue that remains to be addressed concerns the correspondence between prosodic and syntactic structure. The current account requires 2P clitics to be hosted by a *full* prosodic element. In cases like (1), the correspondence of an S-level XP with a *full* prosodic object must be investigated. I argue above that the analysis here assumes a constraint-based phonology for Pashto clauses as in Klein (2000). The constraints responsible for the mapping of syntactic structure to prosodic structure need to be fully fleshed out; a matter I reluctantly leave for future research.

In conclusion, this paper presents an analysis of 2P clitics in Pashto that makes use of word order domains as a means to constrain the linear realisation of clauses containing them. The approach proposed here makes a formal distinction between constraints on hierarchical structure, and those on linear organisation. The obvious benefits of such a strategy are a simplified account of the syntax of 2P clitics, and an account of their placement that makes reference to both syntactic and prosodic information. A derivational account in Roberts (2000) was shown to face certain complications. Specifically, the *clitics-as-agreement* hypothesis does not make obvious gains, as it must be supplemented with a prosodic repair strategy (PI) to account for the basic alternations, and it makes an incorrect empirical prediction regarding clitic doubling. Additionally, the derivational account misses the intuition that the second position effect in Pashto clauses is the result of a prosodic dependency, by dividing the mechanisms that account for 2P clitic placement. The analysis presented here offers a unified treatment of the second position effect, and simplifies the syntactic assumptions required to account for 2P clitic licensing. The domain ordering component of the analysis was shown to follow straightforwardly through the interaction of different constraints, each relatively simple in its formulation. The collaborative effect correctly predicts 2P clitic placement in a variety of different cases, including phrasal hosts (see 1), verbal hosts (see 4) and apparent infixation (see 6). An interesting result of the proposal here concerns the infixation cases and the principle of Lexical Integrity (Bresnan and Mchombo, 1995). At first glance, it appears as though Pashto does violence to the claim that syntactic processes should not make reference to the internal composition of lexical items. However, under the current approach, the syntactic component of the

grammar is not at fault, so to speak. Crucially, I have argued that 2P clitic placement is not a purely syntactic process, but rather one that results from the interaction of syntax, prosody and word order constraints. The account here is strictly compatible with the notion that *syntactic* processes do not reference the internal makeup of words. As a final thought, it is hoped that the research here stimulates further investigation of the role that domain ordering analyses might play in accounting for phenomena that appear to be sensitive to multiple components of the grammar.

Acknowledgments

I would like to thank (in no particular order) Mohammad Dost, Shahab Khan, Andreas Kathol, Geoff Pullum, Judith Aissen, Sandy Chung, Peter Sells, and the audience at TLS 9 for valuable comments and insights. The work presented here is greatly improved as a result of their input. All remaining errors and omissions have been placed deliberately by the author in order to encourage future research in this area.

References

- Babrakzai, Farooq. 1999. *Topics in Pashto Syntax*. Ph.D. thesis, University of Hawai'i at Manoa.
- Bresnan, Joan and Sam Mchombo. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13:181–254.
- Chomsky, Noam. 1995. *The Minimalist Program*, chap. 3, pages 167–217. MIT Press.
- Crysmann, Berthold. 2000. On the placement and morphology of Udi subject agreement. paper presented at the 7th International Conference on Head-driven Phrase Structure Grammar, Berkeley.
- Curry, Haskell B. 1961. Some logical aspects of grammatical structure. In *Structure of Language and its Mathematical Aspects: Proceedings of the Twelfth Symposium in Applied Mathematics*, pages 56–68. American Mathematical Society.
- Dowty, David. 1982. Grammatical relations and Montague grammar. In P. Jacobson and G. K. Pullum, eds., *The Nature of Syntactic Representation*, pages 79–130. Dordrecht: D. Reidel.
- Dowty, David. 1996. Towards a minimalist theory of syntactic structure. In H. Bunt and A. van Horck, eds., *Discontinuous Constituency*, pages 11–62. Berlin/New York: Mouton de Gruyter.
- Ginzburg, Jonathan and Ivan A. Sag. 2001. *Interrogative Investigations: the form, meaning, and use of English Interrogatives*. Stanford, California: CSLI Publications.

- Halpern, Aaron. 1995. *On the Placement and Morphology of Clitics*. Stanford: CSLI.
- Hock, Hans Henrich. 1996. Who's on first? Toward a prosodic account of P2 clitics. In A. L. Halpern and A. M. Zwicky, eds., *Approaching Second: Second Position Clitics and Related Phenomena*, pages 199–270. CSLI Publications.
- Höhle, Tilman. 1986. Der Begriff “Mittelfeld”, Anmerkungen über die Theorie der topologischen Felder. In A. Schoöne, ed., *Kontroversen, alte und neue: Akten des 7. Internationalen Germanisten-Kongresses, Göttingen 1985, Band 3*, pages 329–340. Tübingen: Niemeyer.
- Jaeggli, Osvaldo. 1982. *Topics in Romance Syntax*. Dordrecht: Foris Publications.
- Kathol, Andreas. 1995. *Linearization-Based German Syntax*. Ph.D. thesis, Ohio State University.
- Klein, Ewan. 2000. Prosodic constituency in HPSG. In R. Cann, C. Grover, and P. Miller, eds., *Grammatical Interfaces in HPSG*, pages 169–200. CSLI Publications.
- Kupść, Anna. 2000. *An HPSG Grammar of Polish Clitics*. Ph.D. thesis, Institute of Computer Science, Polish Academy of Sciences and Université Paris 7, Warszawa, Poland.
- Levine, Robert D. and Walt Detmar Meurers. to appear. Head-driven phrase structure grammar: Linguistic approach, formal foundations, and computational realization. In K. Brown, ed., *Encyclopedia of Language and Linguistics*. Oxford: Elsevier, 2nd edn.
- Penn, Gerald. 1999. An RSRL formalization of Serbo-Croatian second position clitic placement. In V. Kordoni, ed., *Tübingen Studies in Head-Driven Phrase Structure Grammar*, no. 132 in Arbeitsberichte des SFB 340, pages 177–197. Tübingen: Universität Tübingen.
- Pollard, Carl J. and Ivan A. Sag. 1987. *Information-based Syntax and Semantics, Vol. 1*. No. 13 in CSLI Lecture Notes. Stanford University: CSLI Publications. Distributed by University of Chicago Press.
- Pollard, Carl J. and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Reape, Mike. 1993. *A Formal Theory of Word Order: A Case Study in West Germanic*. Ph.D. thesis, University of Edinburgh.
- Reape, Mike. 1994. Domain union and word order variation in German. In J. Nerbonne, K. Netter, and C. J. Pollard, eds., *German in Head-Driven Phrase Structure Grammar*, no. 46 in CSLI Lecture Notes, pages 151–197. Stanford University: CSLI Publications.
- Roberts, Taylor. 2000. *Clitics and Agreement*. Ph.D. thesis, Massachusetts Institute of Technology.
- Sag, Ivan A. and Thomas Wasow. 1999. *Syntactic Theory: A Formal Introduction*. Stanford: Center for the Study of Language and Information. Japanese edition (two volumes) – translated and edited by Takao Gunji and Yasunari Harada, appeared in 2001.

- Sportiche, Dominique. 1996. Clitic constructions. In J. Rooryck and L. Zaring, eds., *Phrase Structure and the Lexicon*, pages 213–276. Dordrecht: Kluwer Academic Publishers.
- Tegey, Habibullah. 1977. *The grammar of clitics: evidence from Pashto and other languages*. Ph.D. thesis, University of Illinois at Urbana-Champaign.
- Zwicky, Arnold and Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59:502–513.