Clitics and Agreement

by

Taylor Roberts

B.A. (Spec. Hons.), Linguistics
York University, 1992

M.A., Linguistics
University of British Columbia, 1994

Submitted to the Department of Linguistics and Philosophy
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Linguistics

at the

Massachusetts Institute of Technology

June 2000

Signature of Author: ....
Department of Linguistics and Philosophy
March 10, 2000

Certified by: ....
Shigeru Miyagawa
Professor of Linguistics and Japanese
Thesis Supervisor

Accepted by: ....
Alec Marantz
Professor of Linguistics
Head, Department of Linguistics and Philosophy

Clitics and Agreement

by

Taylor Roberts

Submitted to the Department of Linguistics and Philosophy
on March 10, 2000 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Linguistics

ABSTRACT

A phrase structure is developed for Pashto, the most important Indo-Iranian language for which this task remains to be undertaken. New data show that the placement, ordering, and interpretation of second-position clitics may be derived in the syntax by treating the clitics as agreement heads that identify null arguments in their specifiers. In contrast to previous accounts, the need for phonological operations is drastically reduced, being restricted to sentences containing only a verb (in which prosodic inversion applies as a last resort). In the course of investigating the role of clitics with respect to argument structure and syntactic derivation, several novel phenomena are uncovered that do not exist in better studied languages. Some of the features scrutinized include compound verbs, agreement, aspect, ergativity, word order (scrambling), possessor raising and dislocation, ambiguity, relative clauses, and overt vs. covert movement.

Thesis Supervisor: Shigeru Miyagawa
Title: Professor of Linguistics and Japanese

Acknowledgments

It is a pleasure to acknowledge the many people who have contributed to this dissertation, first among them being my committee members: Shigeru Miyagawa, Ken Hale, and Wayne O'Neil. They have contributed countless ideas and suggestions, and exercised considerable patience with me as I pursued some unworkable dissertation topics. I've been very fortunate to have them as teachers and encouraging advisors.

Thanks to the other professors from whom I have enjoyed taking classes: Noam Chomsky, Kai von Fintel, Suzanne Flynn, Morris Halle, Jim Harris, Irene Heim, Michael Kenstowicz, Jay Keyser, Alec Marantz, David Pesetsky, Roger Schwarzschild, and Ken Wexler.

Thanks to the helpful staff of the department: Anne Cahill, Mary Grenham, Jen Purdy, and Bev Stohl.

I am grateful for the fellowships and grants I have received from MIT, the MIT Working Papers in Linguistics, and the Social Sciences and Humanities Research Council of Canada (752-96-0602). I am extremely fortunate to have had this support, and I do not take it for granted.

This dissertation foolishly attempts to build on the work of Habibullah Tegey. Despite any leaps in understanding that the present work might seem to have made, Tegey's dissertation (1977) remains the best discussion of the most interesting phenomena in Pashto. His reference grammar (Tegey and Robson 1996) has been another essential resource. I was fortunate to have another dissertation on Pashto, by Farooq Babrakzai (1999), arrive in time for me to learn from. I am also grateful to both Tegey and Babrakzai for answering my questions by e-mail.

The Pashtun linguist on whose good will I have depended most has been Jan Mohammad at the University of Arizona. His good nature in answering my many questions, and his skills as a linguist, have been instrumental in my completing the dissertation, as should be obvious from the frequency with which he is cited throughout. Needless to say, it has been very helpful to correspond with a linguist who can judge the grammaticality and ambiguities of my sentences and offer his own sentences in return, complete with *pro* when necessary. This dissertation would not have been as accurate or thorough without his help—although I would also like to absolve Jan of responsibility for any errors I might have made.

I have been greatly encouraged in my research by the many e-mails and telephone calls from Pashtuns and Pathans around the world who have offered their help. Among those consultants who have worked with me in person and whose knowledge is reflected here are Abdul Aziz, Ghulam Isaqzai, Mohammad Mosa Karzai, Nabeela Khatak, Afnan Kundi, and Baryalay Malyar.

Many thanks to Gaurav Mathur for his detailed comments on a draft of the dissertation, and for his last-minute help with proofreading.

The linguistic and Indo-European knowledge of my office mate, Paul Elbourne, was helpful to me on many occasions. Rajesh Bhatt, Simin Karimi, and Barbara Robson have also been helpful with discussions of Pashto, and Elizabeth Cowper's syntax project meetings at the University of Toronto provided pleasant summer venues to present some of this material. And I thank Doug Pulleyblank for, among other things, writing and sharing a Word macro that has made example numbering so easy.

I have been lucky to find many friends here, beginning with my classmates in Ling–94: Christopher Bader, Marie Claude Boivin, Gaurav Mathur, David McKay, Hooi Ling Soh, Luciana Storto, Fleur Veraart, and Susi Wurmbrand. I hope we can gather again someday.

Table of contents

Symbols and abbreviations

x̲	*x* = 2P clitic	LV	light verb
(*x)	ungrammatical if *x* is present	MASC	masculine
(x)	*x* is optional	MOD	modal
*	ungrammatical	NEG	negation
*(x)	ungrammatical if *x* is absent	NEUT	neuter
-	morpheme boundary	NOM	nominative
?	marginally grammatical	NV	non-verbal
{x/y}	either *x* or *y*, but not both	OBL	oblique
1, 2, 3	first, second, third person	PART	participle
2P	second-position	PERF	perfective
ABS	absolutive	PL	plural
ACC	accusative	PN	strong pronoun
ADJ	adjective	POSS	possessive
ADV	adverb	PRES	present tense
ASP	aspect	pro	null pronoun
AUX	auxiliary	RC	relative clause
COMP	complementizer	REFL	reflexive
CONJ	conjunction	SG	singular
CV	compound verb	TRANS	transitive
DAT	dative	VIS	visible
DET	determiner		
DIR	direct case		
EC	empty category		
ERG	ergative		
EXCL	exclamation		
FEM	feminine		
FUT	future		
GEN	genitive		
IMP	imperative		
IMPF	imperfective		
INDIC	indicative		
INTERROG	interrogative		
INTR	intransitive		
INVIS	invisible		

1. Introduction

1.1. Data

This dissertation sketches the basic phrase structure of Pashto, the most typologically interesting and important Indo-Iranian language for which this task remains to be undertaken. While Tegey's (1977) dissertation and his related articles have sparked interest in the language's typologically unusual second-position clitics, most subsequent research pertaining to clitic placement has not improved upon Tegey's analysis. New data show that the placement, ordering, and interpretation of the language's second-position clitics may be derived in the syntax by treating the clitics as agreement heads that identify null arguments in their specifiers. In contrast to previous accounts, the need for phonological operations is drastically reduced, being restricted to sentences containing only a verb in addition to the clitic(s) (in which prosodic inversion applies as a last resort). Pashto's second-position clitics are thereby shown not to be as unusual as has hitherto been believed—a conclusion that should not be surprising within a framework like the Minimalist Program (Chomsky 1995), which hypothesizes that an invariant Universal Grammar underlies all languages, despite surface appearances to the contrary. Chapters 1 and 2 explore the general phrase structure of the language, with a particular focus on properties of agreement and constituency of compound verbs. This background sets the stage for a detailed discussion of second-position clitics in chapters 3 and 4.

Pashto has approximately twenty million speakers in Afghanistan and Pakistan. This introductory chapter describes some basic properties of the language: its phonology (§1.2), word order (§1.3), nominal morphology (§1.4), and verbal morphology and agreement patterns (§1.5).[1] The presentation is heavily indebted to Tegey and Robson's (1996) morphological classifications, while the romanized orthography is modeled closely upon that of Penzl (1955: 14–37). As data from several varieties of Pashto are presented, it will be useful to identify the source of each

---

[1] For additional information on the language, such works as Penzl (1955), Shafeev (1964), MacKenzie (1987), Skjærvø (1989), Tegey and Robson (1996), Elfenbein (1997), and Babrakzai (1999) may be consulted.

sentence, in order to minimize apparent inconsistencies between examples that are due to dialectal variation. Sentences from published sources are cited in the conventional way. Data gathered during fieldwork are cited by a place name after the example (e.g., 'Kandahar' or 'Laghman'), indicating the consultant's place of origin. The only exceptions are data that have been contributed by Pashtun linguists, which are cited as personal communications ('p.c.'); of these, Habibullah Tegey is from Kabul, and Farooq Babrakzai is from Zadran (Jadran) in Paktia; Jan Mohammad was raised in Paktia, and speaks Eastern Afghanistan Pashto.

1.2. Phonology and orthography

Inventories of consonant and vowel phonemes are presented in (1–2). These charts also serve as a key to the orthography.

(1)

	labial		dental	alveolar	palatal		retroflex	velar		glottal	
stops	p	b	t	d			tt	dd	k	g	q[2] '[3]
fricatives	f			s	z	sh	zh	ss	zz	kh	gh	h	qh[4]
nasals		m	n					nn			
liquids				l	r			rr			
glides		w				y					

(2)

	front	mid	back
high	i		u
mid	ee	e	o[5]
low	a		aa

[2] Voiceless pharyngealized velar stop.
[3] Glottal stop.
[4] Voiceless pharyngeal fricative, which Penzl (1955) symbolizes with underlined <h>.
[5] Penzl (1955) uses <oo>.

Pashto is unique among Iranian languages in having a series of retroflex consonants, shared with neighboring but more distantly related languages like Urdu. The retroflex nasal /nn/ and liquid /rr/ have a ballistic, flap articulation, especially before vowels; in other environments, they are more like approximants. The retroflex fricatives are pronounced in the southern variety often called 'soft Pashto' (medial *sh* in the language name 'Pashto' representing a voiceless retroflex fricative), but have merged with the corresponding voiced and voiceless velar fricatives in more northerly ('hard Pashto') areas. This variation accounts for the name of northern varieties of the language often being rendered 'Pakhto'.

The phonemes /f q qh '/ occur in words borrowed from Arabic and Farsi, but are pronounced only in formal, educated speech. In informal speech, /f/ is pronounced [p], /q/ and /qh/ are deleted or pronounced as [k], and /'/ is deleted.

The mid vowels /ee/ and /o/ participate in a vowel harmony process, raising to [i] and [u] respectively when the following syllable contains a high vowel. The other mid vowel, schwa /e/, has more elusive behavior (as it does in many languages), and with respect to vowel harmony may surface variously as [i] or [u], depending also on the rounding of surrounding consonants.

Each word bears a primary stress, which is determined idiosyncratically for each lexical item. Unless indicated otherwise by an acute accent, primary stress falls on the final syllable of a cited form (the most common location). Becka (1969) provides a good description of stress patterns, but the system awaits a fuller treatment (which surely will also help to inform an analysis of the specifically prosodic properties of clitic placement, not treated in depth here).

Despite having noteworthy phonetic characteristics and intriguing patterns of vowel harmony and coalescence, there has been little phonological analysis of Pashto other than Kaisse (1981) and parts of Tegey (1977). The only acoustic study is Bell and Saka (1983), which examines Pashto's distinctive reversed-sonority consonant clusters in such words as *wradz* 'day', *lmar* 'sun', and *wlaarr* 'upright'.

1.3. Word order

Pashto is fairly rigidly head-final across lexical categories, while several functional categories are head-initial—a split that will emerge as the facts of the language are presented in this chapter and the next. The basic word order is SOV, as illustrated below. (The structure of compound verbs— such as *newishta key* 'write do', below—will be discussed at length in Chapter 2.)

(3)   Sur Gwel   kitab   newishta   key
        Sur Gwel   book   write   do
        'Sur Gwel wrote a book' (Kandahar)

For some speakers, the order OSV, as in (4), is unnatural unless there is a heavy pause after the direct object, and a context (for example, in answer to the question, 'Who wrote a book?').

(4)   *kitab   Sur Gwel   newishta   key
        book   Sur Gwel   write   do
        'Sur Gwel wrote a book' (Kandahar)

Many speakers, however, allow a freer word order (contra the statement of MacKenzie 1987: 563), especially when the grammatical roles of the referents denoted by the nominals are established by context, or when the nominals bear case-markings that disambiguate their grammatical functions (Tegey 1979: 379; Babrakzai 1999:61), as frequently happens in past tense sentences, which show ergative case and agreement. (The case system is explained in the next section.) The following sentences illustrate SOV and OSV order respectively, and are equivalent with respect to their meaning and prosody (Jan Mohammad, p.c.):

(5)　　　a.　spi　　　　　　pisho　　　　khog　krr-a

　　　　　　　dog(OBL MASC)　　*cat(DIR FEM)*　　*hurt*　*do(PAST PERF)-FEM3SG*

　　　　　　　SOV: 'the dog hurt the cat'

　　　　　b.　pisho　　　　　spi　　　　　khog　krr-a

　　　　　　　cat(DIR FEM)　　*dog(OBL MASC)*　*hurt*　*do(PAST PERF)-FEM3SG*

　　　　　　　OSV: 'the dog hurt the cat' (Jan Mohammad, p.c.)

This freedom of word order is not possible in present tense, since the two arguments receive the same case-marking. Compare (5a) with (6), in which the OSV interpretation is ungrammatical:

(6)　　　　　spay　　　　pisho　　　　khog-aw-i

　　　　　　　dog(DIR)　　*cat(DIR)*　　*hurt-TRANS(PRES IMPF)-3SG*

　　　　　　　SOV: 'the dog is hurting the cat' / OSV: *'the cat is hurting the dog'

　　　　　　　　　　　　　　　　　　　　　　　　　　　　　(Jan Mohammad, p.c.)

The OSV interpretation becomes available by following the fronted object with a heavy pause (topicalization), as indicated in (7) by the comma:

(7)　　　　　spay,　　　　pisho　　　　khog-aw-i

　　　　　　　dog(DIR)　　*cat(DIR)*　　*hurt-TRANS(PRES IMPF)-3SG*

　　　　　　　'the cat is hurting the dog' (Jan Mohammad, p.c.)

The OSV interpretation may also be obtained by left-dislocation, as illustrated in (8) below. Left-dislocation structures contain a clitic that corefers with the left-dislocated nominal. Below, this clitic is 3sg accusative *yee*. Here and throughout, 2P clitics are underlined:

(8)   spay,   pisho   <u>yee</u>   khog-aw-i

dog(DIR)   cat(DIR)   3SG   hurt-TRANS(PRES IMPF)-3SG

'the dog, the cat is hurting him' (Jan Mohammad, p.c.)

The availability of both topicalization (7) and left-dislocation (8) will be seen in chapter 3 to be a useful tool in determining why relative clauses appear (oddly) to induce clitic-doubling.

 Word order is similarly flexible in sentences having ditransitive verbs—again, as long as the grammatical functions of the arguments are clear from context or case-marking. There is thus no need for pauses or special intonation in any of the following sentences:

(9) a. [$_S$ Khálid] [$_{IO}$ Ásiye ta] [$_{DO}$ kitáb] [$_V$ weer krro]

   Khalid   Asia to   book give do(PAST PERF)

   (OBL)   (OBL)   (DIR)

 b. [$_S$ Khálid] [$_{DO}$ kitáb] [$_{IO}$ Ásiye ta] [$_V$ weer krro]

 c. [$_{DO}$ kitáb] [$_S$ Khálid] [$_{IO}$ Ásiye ta] [$_V$ weer krro]

 d. [$_{DO}$ kitáb] [$_{IO}$ Ásiye ta] [$_S$ Khálid] [$_V$ weer krro]

 e. [$_{IO}$ Ásiye ta] [$_S$ Khálid] [$_{DO}$ kitáb] [$_V$ weer krro]

 f. [$_{IO}$ Ásiye ta] [$_{DO}$ kitáb] [$_S$ Khálid] [$_V$ weer krro]

   'Khalid gave the book to Asia' (Boraka)

As illustrated above, all orderings of the arguments are acceptable, as long as the verb appears finally.[6]

 Numerals and adjectives precede any nouns they modify, suggesting that the lexical category NP is head-final:

---

[6] A similar paradigm is given by Babrakzai (1999: 14).

(10)   [NP   pindze   khkwelee   péeghlee]

              five   *pretty*   *girls*   (Tegey and Robson 1996: 85)

Possessive phrases (marked by the preposition *dee/de*) and other PPs also precede the nouns they modify:

(11)   a.   de   dzhan   plaar

              POSS   *John*   *father*

              'John's father' (Babrakzai 1999: 31)

         b.   de   kitaab   zzaanee   de   mudir   de   wror   kor

              POSS   *book*   *house*   *POSS*   *manager* *POSS*   *brother*   *house*

              'the house of the brother of the manager of the library' (Babrakzai 1999: 31)

         c.   dee   maktab   dee   shaagerdáano   dee   dars   dee   dawree

              POSS   *school*   *POSS*   *students*   *POSS*   *study*   *POSS*   *period*

              dee   wakht   le   khwakhey   na   ddáka   yaaduna

              POSS   *time*   *from*   *happiness*   *from*   *full*   *memories*

              'memories full of happiness of the time period of students' studying at school'

                                                                      (Tegey and Robson 1996: 172)

On the other hand, DP (a functional category), is head-initial, as may be seen from the following sentence, in which the determiner *daa* appears initially:

(12)   [DP   daa   peeghla]   kitabúna   khe   lwal-i

              that   *girl*   *books*   *well*   *read-3SG*

              'That girl reads books well' (Tegey and Robson 1996: 88)

14

With regard to the PP projection, the language appears to exhibit mixed headedness. There is a single postposition, *ta* 'to'—exemplified in (9) above—and two prepositions: *dee/de*[7], which marks possessive NPs—illustrated in (11) above—and *pe* 'by means of, with, at'. The remaining members of the PP category are ambipositions, an example of which appears in (11c) above: *le ... na* 'from'. Ambipositions are discussed more fully in section 2.5; a complete list of them may be found there in (78).

Subordinate clauses are preceded by the complementizer *tshee/tshi*[8] (which is usually obligatory), and so CP, like several other functional categories, is head-initial:

(13) Mamaad fiker kewi [CP **tshi** de[9] Sur Gwel day khwaass dey]

     Mamaad *thought* *do* *COMP* *POSS* *Sur Gwel* *him* *like* *be*

     'Mamaad thinks Sur Gwel likes him' (Kandahar)

(14) a. de Sangin khyal dee [**tshi** mashem uda dee]

     POSS *Sangin* *thought* *be* *COMP* *baby* *sleep* *be*

     'Sangin thinks the baby is sleeping' (Yusufzai)

    b. *de Sangin [**tshi** mashem uda dee] khyal dee[10]

     POSS *Sangin* *COMP* *baby* *sleep* *be* *thought* *be*

---

[7] The two forms of the possessive preposition (*dee* vs. *de*) appear to be in free variation, although it is not unlikely that they participate in the process of vowel harmony described in sec. 1.2.

[8] See the previous note. The two forms of the complementizer (*tshee* vs. *tshi*) similarly seem to be in free variation, although vowel harmony should not be excluded as a factor influencing the selection of these forms.

[9] The embedded subject appears in possessive form (*de Sur Gwel*) because 'like' is a psych-predicate; see the next section for discussion.

[10] See previous note. The subject is not the possessor of *khyal* 'thought' (e.g., "Sangin's thought is that...."). Rather, the possessive form of the subject is due to the verb being a psych-predicate, which is compound (*khyal dee*).

(15)   kéla   [CP **tshi**   dey   kor   te   raghey]   Sur Gwel   bide   su[11]

           when   COMP   *he*   *home*   *to*   *arrive*   *Sur Gwel*   *sleep*   *be*

           'When he arrived home, Sur Gwel fell asleep' (Kandahar)

Complex sentences of the kind in (13–14), which are selected by a matrix verb like 'think', are the only structures that permit the verb to appear non-finally; indeed, as the ungrammaticality of (14b) indicates, such complement clauses *must* follow the verb that selects them. In contrast, relative clauses are positioned more freely with respect to the nominal head that selects them (though note again that CP is invariably head-initial):

(16)   a.   hagheey   hagha   mayshem   [CP **tshi**   uda   dee]   khkol   krro

             she   DET   *baby*   COMP   *sleep*   *be*   *kiss*   *do*

            b.   hagheey   hagha   mayshem   khkol   krro   [CP **tshi**   uda   dee]

             she   DET   *baby*   *kiss*   *do*   COMP   *sleep*   *be*

            'she kissed the baby who is sleeping' (Yusufzai)

Hindi also has a head-initial CP, despite the otherwise head-final character of the language. Anoop Mahajan (p.c.) has entertained the idea of a head-final CP for Hindi, which might be possible if the complementizer of a subordinate clause actually belonged to the superordinate clause. Unfortunately, there is little evidence for positing a head-final CP, other than in the complementizer *dzeka tshee* 'because', which may occasionally be discontinuous, its first element preceding the matrix verb:[12]

---

[11] The *wh*-word *kéla* 'when' is shown outside CP (rather than in Spec/CP), since Pashto does not have *wh*-movement. Section 3.6 will show that seemingly fronted *wh*-words are actually left-dislocated (i.e., merged in their surface position, rather than moved there).

[12] This sentence is from Mir Mindi Shah Mindi, 'The Grave' (Tegey and Robson 1996: 214).

(17)	daa	tapus	<u>mee</u>	**dzéka**	wekrre	[	**tshee**	hagha	ba
	this	*question*	*1SG*	*because*	*did*		*COMP*	*3SG*	*would*
	dee	qaber	na	dder	zyaat	weereedo]			
	from	*grave*	*from*	*very*	*heavy*	*fear*			

'I asked this question because he was terrified of graves'

Yet even in (17), if CP were head-final, *dzeka* 'because' would be expected to follow the matrix auxiliary verb *wekrre* 'did', since CP is the highest projection. Moreover, there would be no explanation for why the complementizer appears initially in all subordinate clauses, irrespective of their position with respect to the matrix verb, as in sentences like (15–16) above. Because of these difficulties, a head-initial CP will be assumed.

The structure of the clause begins to take the following form:

(18)
```
        CP
         \
          C'
         /  \
        C    TP
            /  \
       SUBJECT  VP
              /  \
     COMPLEMENT   V
```
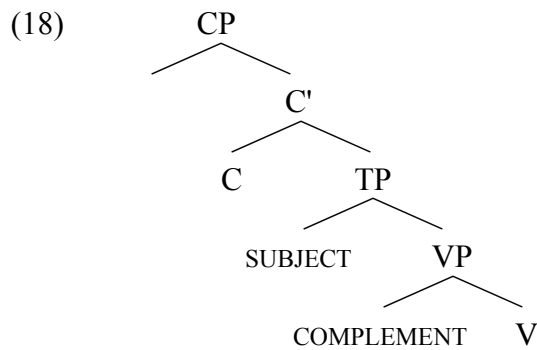
It remains to flesh out the functional categories lower than CP, which will be possible after examining agreement within the clause, the subject of the following chapter.

1.4. Nominal morphology and case

In a language like English, a single case (nominative) is strongly related to derived subject position, Spec/TP. This correlation between a single case and the surface subject position is

disrupted in Pashto, for two reasons. First, the language has a pattern of split-ergativity similar to that in Hindi/Urdu, except that Pashto defines the split on tense, rather than on aspect (though this distinction will be refined in the following chapter); see (34) below for an example. Second, subjects may take the form of possessive or dative phrases, depending on the predicates that select them, as well as on the degree of volition attributed to the referent. Recognizing that the case of an NP is not always correlated with its structural position will be useful in the analysis to be proposed in chapter 3, where it will be shown that the positioning of second-position clitics in the clause may be handled straightforwardly by permitting non-nominative NPs to occupy Spec/TP.

Nouns bear features of gender (masculine and feminine), number (singular and plural), and case (direct and oblique). Gender of nouns is reflected by the varying forms of the verbs and adjectives that agree with them. A noun such as *patlun* 'pants' may therefore take a variety of forms, depending on its number and grammatical role (Tegey and Robson 1996: 50):

(19)

	sg	pl
direct	patlun	patlanúna
oblique	patlaane	patlanúno

The four-way distinction exemplified by (19) is not made by all nouns; many nouns do not distinguish direct and oblique singular forms, for example. The class membership of nouns is determined largely arbitrarily, and there is little consensus on how to divide them. Tegey and Robson (1996) posit four classes of masculine nouns and three classes of feminine, while recognizing a good number of irregular forms.

The two cases encode a variety of grammatical functions and, as mentioned above, display an ergative pattern in past tense. In the chart below, 'subject' is meant to refers to subjects of transitive and unergative verbs only, since subjects of unaccusative verbs behave as objects:

18

(20)

	direct	oblique
present	subject; object	object of adposition[13]
past	object	subject; object of adposition

Some members of the set of singular strong pronouns initially appear to show an additional case distinction, differentiating direct objects from objects of adpositions, as shown in the paradigm in (21) (adapted from Tegey and Robson 1996: 69). As will be explained below, however, 'accusative' is not a distinct case; the term is intended merely to identify the direct object in a present tense sentence, in order to show how case-assignment is split according to person.

(21)   Strong pronouns (singular)

			direct	accusative	oblique (obj. of P)
1sg			ze	maa	maa
2sg			te	taa	taa
3sg	vis	masc	day	day	de
		fem	daa	daa	dee
	invis	masc	agha	agha	aghe
		fem	agha	agha	aghee

The identity of form of the pronouns within double-lined boxes suggests that—like full NPs— strong pronouns likewise only show two-case distinctions, direct and oblique. Third-person pronouns are like full NPs (which are also, of course, third-person) in receiving direct case when they are the direct object of a present tense sentence. In contrast, first- and second-person

---

[13] The term *adposition* refers as a group to prepositions, postpositions, and ambipositions.

pronouns, when they are objects, receive oblique case in present tense.[14] As shown below, the so-called accusative 1sg and 2sg pronouns in (21) are restricted to direct object position in present tense:

(22) a. ze   taa/*te   daftar   ta   leeg-em

         PN1SG(DIR)   PN2SG(ACC)/PN2SG(DIR)   *office   to   send-1SG*

         'I am sending you to the office'

     b. te   maa   daftar   ta   leeg-ee

         PN2SG(DIR)   PN1SG(ACC)   office   to   send-2SG

         'you are sending me to the office' (Babrakzai 1999: 60)

Direct case is used for subjects and objects in present tense, and for objects in past tense. And, also like full NPs, pronominal subjects appear in oblique (ergative) case in past tense:

(23) a. minee   ze   pe   baagh   kee   we   lid-em

         Meena(OBL)   PN1SG(DIR)   at   garden   in   PERF   saw-1SG

         'Meena saw me in the garden'

     b. maa   mina   pe   baagh   kee   we   lid-a

         PN1SG(OBL)   Meena(DIR)   at   garden   in   PERF   saw-FEM3SG

         'I saw Meena in the garden' (Babrakzai 1999: 61)

If the person-split is borne in mind, the chart in (21) may therefore be simplified as follows:

---

[14] The split between first- and second-person nominals vs. third-person nominals may be defined simply as being between discourse participants and non-participants (a split familiar from many languages); see section 4.2 for evidence from clitic ordering and interpretation that the person-split is structurally represented.

(24) Strong pronouns (singular) — simplified

			direct	oblique
1sg			ze	maa
2sg			te	taa
3sg	vis	masc	day	de
		fem	daa	dee
	invis	masc	agha	aghe
		fem	agha	aghee

Whereas singular strong pronouns bear two cases, plural strong pronouns have a single form, regardless of their function in a sentence:

(25) Strong pronouns (plural)

		all functions
1pl		mung
2pl		táasee
3pl	vis	duy
	invis	aghuy

Strong pronouns appear in the same positions as full NPs (i.e., in unmarked SOV order), as illustrated by the above sentences. A strong pronoun is used when its referent is emphasized; discourse-neutral (topic) pronouns take the form of second-position clitics, to be discussed in chapter 3.

21

As was mentioned at the beginning of this section, oblique-marked NPs may fill subject position. Psych-predicates (Belletti and Rizzi 1988) such as 'like' (13) and 'think' (14), select a possessive subject, which comprises the possessive preposition *dee/de* and its oblique-marked complement; other predicates having this property are 'dislike', 'feel hot/cold', and 'have a fever' (Tegey and Robson 1996: 184–188). Still other predicates require their subjects to appear as complements of locative, dative, or ablative adpositions, although the subject NPs themselves still appear in oblique case (Babrakzai 1999: ch. 7). Moreover, the subject's degree of volition may often be indicated by these varying ways of marking the subject. Consider the following sentences:

(26)   a.   **de   laylaa** delta pinze kaala   teer   shw-el

               POSS   Layla   here   five   years(DIR) passed became-MASC3PL

     b.   **pe   laylaa baandee** delta pinze kaala   teer   shw-el

               LOC   Layla   on   here   five   years(DIR) passed became-MASC3PL

     c.   **laylaa**   delta pinze kaala   teer   krr-el

               Layla(OBL) here five   years(DIR) passed did-MASC3PL

               'Layla spent five years here' (Babrakzai: 179–180)

The subject, 'Layla', receives the same case-marking (oblique) in each of these sentences, all of which have the same literal meaning, although the subject appears within a PP in the first two sentences. The possessive-marked subject in (26a) receives the most neutral interpretation of these three sentences. In (26b), the subject is flanked by the locative ambiposition *pe ... baandee*, which suggests that Layla had no choice in her stay. In contrast, the verb in (26c) is transitive, and the subject is a bare NP, though still marked oblique (because of past-tense ergativity); in this

sentence, Layla is interpreted as having exercised volition in her stay, deliberately spending five years in one place.

Another area of sentence formation in which the case of an NP is not correlated to its structural position concerns unergative verbs. It was mentioned with respect to the chart in (20) that subjects of unaccusative verbs behave as objects, in that they receive direct case in both present and past tense. A small class of intransitive verbs, however, select an oblique-marked subject NP in past tense, exactly as if the verb were transitive. This case-marking is illustrated in (27) with the unergative verb *khandel* 'laugh'; see also (169) below for other variants within this paradigm.

(27) a. kháand-em

  laugh-1SG

  'I am laughing'

  b. maa khand-él(e)

  PN1SG(OBL) laugh-MASC3PL(PAST)

  'I was laughing'

  c. khand-él(e) mee

  laugh-MASC3PL(PAST) 1SG

  'I was laughing' (Tegey and Robson 1996: 188)

The 1sg subject in present tense (27a) is identified by the verbal agreement suffix. In past tense, however, the subject appears in oblique case, either in the form of a strong pronoun, as in (27b), or as a clitic, as in (27c). In past tense, the verbal agreement suffix does not agree with the subject, but rather is the default, 3pl masculine. Other unergative verbs include 'bray', 'whinny', 'cry', 'sneeze', 'cough/spit', 'roar', 'dance', 'swing', 'swim', 'jump', and 'bark'.

1.5. Verbal morphology and agreement

Verbs show both subject and object agreement. Understanding agreement and its position in the clause is crucial to understanding Pashto's second-position clitics. A more detailed examination of verbal agreement is undertaken in chapter 2, but the basic verbal morphology is outlined below. Agreement in the language is wryly described by MacKenzie (1987: 564): '... there is little to be said except that, where the forms permit it, it is all-pervading.'

1.5.1. Simple verbs

The morphological structure of verbs may be represented schematically as follows, with the symbol # indicating positions in which a second-position clitic potentially may appear, if the sentence has no other suitable host for the clitic—a situation that can arise when the verb licenses *pro*-drop:

(28)   [$_V$ # aspect # negation # stem – agreement #]   (Babrakzai 1999: 51)

Verbs have different forms depending on their tense (past vs. present) and aspect (perfective vs. imperfective). Perfective is productively marked by the stressed proclitic *we* (the vowel of which is rounded when followed by a labial consonant), while past tense is marked (on regular verbs) by the stressed suffix -*el*, which is also the infinitive marker. Thus, the four possible tense/aspect stems for the simple verb *tarrel* 'to tie' are as follows (Tegey and Robson 1996: 99):

(29)

	present	past
impf	tarr-	tarr-él-
perf	wé-tarr-	wé-tarr-el-

Many intransitive verbs bear the suffix *-eeg* in present tense, and *-eed(él)* in past tense. A common transitive/causative suffix is *-aw*. These suffixes are reduced forms of the intransitive and transitive auxiliaries, for which see (38–39) below. The final element of any verb is one of the suffixes from the paradigm in (30), which agrees with relevant arguments (depending on tense and aspect) in person and number. The third-person suffixes also show gender agreement in past tense.

(30)   Verbal suffixes

			sg	pl
1			-em	-u
2			-ee	-ey
3	present		-i	-i
	past	masc	-e	-e
		fem	-a	-ee

These suffixes license *pro*-drop in subject position in present tense, and in object position in past tense, as will be amply illustrated in chapter 3.

The perfective morpheme *we* has been inconsistently regarded (even within single works) as a prefix or as a free morpheme. This morpheme is not actually a prefix, as it is separable from the verb stem. The fact that it appears in the vicinity of the verb with other functional morphemes—with this derived verbal complex usually bearing a single primary stress—suggests that, morphophonologically, the perfective morpheme *we* is a proclitic. In (31a), the verbal sequence *we ne khwaarra* behaves as a single word, bearing stress on the initial syllable; when the object is dropped, and the subject changed into the second-position ergative clitic *yee*, however, as in (31b), perfective *we* is divided from the verb stem by the second-position clitic (there being no other suitable host for the clitic in the sentence):

(31)   a.   ahmad   manna   we   ne   khwaarr-a

Ahmad   apple(FEM)   PERF   NEG   ate-FEM3SG

'Ahmad did not eat the apple'

   b.   we   yee   ne   khwaarr-a

PERF   3SG   NEG   ate-FEM3SG

'he did not eat it' (Babrakzai 1999: 51)

The negative morpheme *ne* exhibits identical behavior. It has sometimes been treated as a prefix, because it forms a single stress-bearing word with the verb stem. However, if the perfective morpheme is dropped from (31b), the negative morpheme is also divisible from the verb stem by the second-position clitic *yee*:

(32)   ne   yee   khwaarr-a

NEG   3SG   ate-FEM3SG

'he was not eating it' (Babrakzai 1999: 52)

Because of this behavior, perfective *we* and negative *ne* are regarded here as proclitics (to the verb), rather than as prefixes. With respect to clitic placement, they contrast strikingly with the verbal agreement suffixes, which are true affixes, and may never be divided from their verb stem:

(33)   a.   *khwaarr-yee-a

ate-3SG-FEM3SG

b. khwaarr-a   yee

    ate-FEM3SG   3SG

    'he was eating it' (Babrakzai 1999: 52)

The morphological distinction between proclitics and affixes will be handled by assuming that proclitics (perfective *we* and negative *ne*) occupy syntactic projections distinct from the verb (AspectP and NegP respectively). A late morphophonological rule identifies these morphemes as proclitics and merges them with the verb to form a single prosodic word at PF. Verbal stems and their suffixes (which include past tense *-el* and the agreement suffixes, in that order) are assumed to enter the derivation as a single, fully formed word (Chomsky 1995), which is why clitics may not intervene between these morphemes. (Similar affixal asymmetries occur in English; see (299) below, for example.[15]) This assumption also entails the necessary result that verbal agreement affixes may co-occur with overt arguments, in contrast to clitics, which can have different behavior in this regard (Jaeggli 1982: 55 n. 10).

These assumptions also have implications for the analysis of verb movement to the Tense node. If the verb moves to T overtly, and if TP is head-final, then verb movement will generally not be detectable, since Pashto is verb-final. On the other hand, if TP is head-initial (like several other functional categories), then the verb must *not* move at overt syntax, since the verb must remain final at Spell-Out. If the verb does move to T after Spell-Out, then TP may be either head-initial or head-final, without any consequences for surface word order. The choice between these alternatives (overt movement to head-final TP vs. covert movement to either head-initial or head-final TP) does not bear crucially on the analyses to be presented throughout, and so the issue will not be considered in more depth than empirical considerations warrant.[16]

---

[15] Another Iranian language, Kurdish, is unlike Pashto in permitting the second-position ergative clitic to intervene between the verb stem and the object-agreement suffix (VanLoon 1997: 162, 166).

[16] But see sec. 3.4.2.2 for evidence that the verb does not move to T in overt syntax.

As mentioned in the previous section, past tense sentences are inflected on an ergative/absolute pattern:

(34)   a.   sarr-**ay**   mann-a   khwr-**i**

             man(MASC)-DIR SG   *apple(FEM)-DIR SG*   *eat-PRES3SG*

             'the man is eating the apple'

        b.   sarr-**i**   mann-a   khwar-él-**a**

             man(MASC)-OBL SG   *apple(FEM)-DIR SG*   *eat-PAST-FEM3SG*

             'the man was eating the apple' (Tegey and Robson 1996: 182)

Both sentences have the same form of the direct object, in the unmarked direct case. The subject in present tense (34a) is also in direct case, resulting in a fairly rigid SOV order for some speakers, since the subject and object are not morphologically disambiguated. The ergative subject in (34b), however, appears in the marked, oblique case. The form of the verb also changes in these sentences, agreeing with the subject in (34a), but with the object in (34b): the classic ergative 'split'. In sentences with simple verbs, case and agreement are therefore correlated. For speakers having a rigid word order, past tense sentences are also SOV, as in (34b), mirroring the word order of present tense clauses. Many speakers have a freer word order, however, as was mentioned above.

Ergativity in Pashto, as in Hindi, is primarily morphological, rather than syntactic, suggesting that subjects of both present and past tense sentences appear in the same position, despite their different ergative and nominative markings. Following a familiar approach to ergativity, it will be assumed that past tense verbs do not assign case to their objects, perhaps because they are related historically to the passive (Mahajan 1990, Bittner and Hale 1996, Murasugi 1997). Lacking case, the object must move out of VP and into Spec/AgrOP in order to receive structural case, perhaps from T (which assigns the unmarked, direct case); the subject

28

receives inherent case, which is ergative. Nevertheless, as will become clear in chapter 3, all NPs (whether the sentence is present or past tense) escape VP, either to satisfy the EPP, or to erase agreement features. At Spell Out, then, the structure of a clause having an overt subject and object is minimally and approximately as follows:

(35)

$$
\begin{array}{c}
\text{TP} \\
\diagup\diagdown \\
\text{SUBJECT}_i \quad \text{AgrO} \\
\diagup\diagdown \\
\text{OBJECT}_k \quad v\text{P} \\
\diagup\diagdown \\
t_i \quad \text{VP} \\
\diagup\diagdown \\
t_k \quad \text{V}
\end{array}
$$
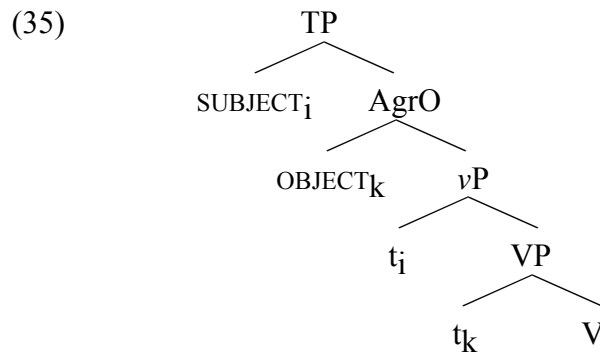
The need to delete EPP and/or agreement features (rather than to receive case) is assumed to be the principal force driving NP movement; the assignment of case is incidental, except for dative arguments, which must be followed by a case-assigning postposition.

1.5.2. Auxiliaries

The most common verbs tend to have irregular past and present stems, and sometimes different stems specifically for third-person masculine forms; see Tegey and Robson (103–104) for a list of such verbs. Among such suppletive verbs is the one meaning 'to be', which does not have an infinitive form. The full forms are given in (36–37), which can be seen to contain the regular agreement suffixes from (30):

(36)　Present tense forms of 'to be'

		imperfective		perfective	
		sg	pl	sg	pl
1		yem	yu	sem	su
2		yee	yáastey	see	sey
3	masc	da/dey/day	di	si	si
	fem	da			

(37)　Past tense forms of 'to be'

		imperfective		perfective	
		sg	pl	sg	pl
1		wem	wu	swem	swu
2		wee	wey	swee	swey
3	masc	we	we	swe	swél(e)
	fem	wa	wee	swa	

The stems of intransitive and transitive auxiliaries, to which the agreement suffixes in (30) are added, are given below (Tegey 1977: 95). Although they are most commonly used to form compound verbs (the subject of the next chapter), they may also be used as main verbs with the respective meanings 'to become' and 'to do, to make'.

(38)　Intransitive auxiliary

	present	past
imperfective	(k)éeg-	(k)eed-
perfective	s-	sw-

(39)   Transitive auxiliary

	present	past
imperfective	(k)aw-	(k)aw(él)
perfective	k-	krr-

Some forms of 'to be' and the auxiliaries have optional variants containing the past tense suffix *-el*, which appears as the penultimate syllable in forms that do not already contain *-el* (Tegey and Robson 1996: 96–98).

These somewhat idealized paradigms are based on the 'central' dialect of Habibullah Tegey, but in actual speech there is considerable variation, especially for vowels, both in individual speakers (due to vowel harmony and rate of speech), and in different dialects. The same is true of all languages, of course, and so Pashto is not unusual in this regard.

1.6. Outline

The facts presented until now have suggested that only tense is relevant to ergativity. If that were the case, Pashto would appear to be unlike its better studied Indo-Iranian sister, Hindi/Urdu, which instead has aspect-conditioned ergativity (nominative/accusative in imperfective, and ergative/absolutive in perfective). In chapter 2, a detailed examination of asymmetries in clitic placement and agreement patterns in compound verbs will reveal that Pashto does indeed evince asymmetries that are crucially driven by aspect. It will also be shown that aspect determines whether the constituents of a compound verb form one unit or two, which is important for understanding clitic placement. Moreover, the most explanatory arrangement of clause structure suggests that, although Pashto appears to be head-final, it is specifically only the lexical projections (VP, NP) that are head-final.

The information about aspect and constituency detailed in Chapter 2 will be helpful in understanding the material in Chapter 3, which examines second-position (2P) clitics. More than any other language, the 2P clitics of Pashto have been particularly recalcitrant to satisfactory analysis. The 2P pronominal clitics will be shown to be agreement morphemes (in contrast to previous analyses, which have regarded them as arguments). By articulating the positions of agreement projections and examining some properties of scrambling, this chapter will show that the placement of 2P clitics does not require syntactic movement of the clitics, but that their placement is due to their being merged late in the derivation, above VP.

Chapter 4 examines the order of clitics within the 2P cluster, and shows that their ordering may similarly be explained by merging clitics directly into their surface positions, without any need for a morphological template. The overall analysis has a simplicity that is unprecedented in discussions of Pashto, as it shows that most 2P clitic phenomena are due to general syntactic principles, with only a small residue of cases requiring prosodic inversion—itself a principled, last resort mechanism of the phonological component. The need for a 'second position' in either the syntax or the phonology will therefore be eliminated.

2. Complex verbs

2.1. Aspect-driven asymmetries

Adjectives and nouns combine with transitive and intransitive auxiliaries to form compound verbs, which are a fairly open class, and constitute the majority of verbs in the language. The auxiliary verb always appears at the end of the sentence, bearing the agreement suffixes listed in (30) in chapter 1. Several examples of compound verbs have already been seen in the previous chapter, such as *fíker kewi* 'thought do' (13) and *khkol krro* 'kiss do' (16). Although complex predicates in Indo-Iranian languages have received some attention (see Ramchand 1991 for Bangla, Butt 1995 for Urdu, and Karimi 1997a,b for Persian), such predicates in Pashto have different properties that offer insight into the language's unusual patterns of agreement and clitic placement.

Compound verbs show that aspect plays a crucial role in determining syntactic and morphological constituency in Pashto. In particular, compound verbs in perfective aspect behave as two units rather than one. It will be suggested that the best way to account for these aspect-driven asymmetries is to project aspect features between VP and TP (following similar representations in Hendrick 1991, Ramchand 1997, 1998, Cinque 1999, and Iatridou et al. to appear):

(40)
```
                    TP
                    |
                  AspP
                 /     \
             VP          Asp
            /  \          |
        kiss do      [PERF/IMPF]
```
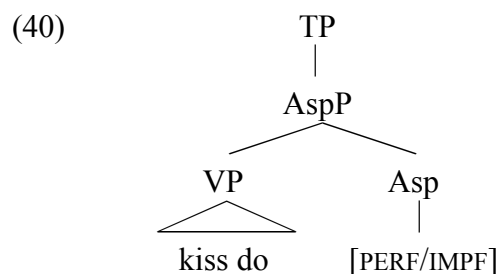
Pashto's aspect-driven asymmetries may be explained by regarding perfective aspect as a strong feature, which entails that a perfective auxiliary verb will move to Asp before Spell Out (i.e., in overt syntax) in order to erase the feature. Because the verb remains sentence-final after Spell

Out, AspectP is regarded as head-final. In contrast, imperfective aspect is a weak feature, which by Procrastinate entails that the verb will not move until after Spell Out (i.e., it will erase the feature covertly, at LF). The verb enters the derivation fully inflected from the lexicon and remains head-final, within VP, at Spell Out (see section 1.5.1 for details). There is no need for overt movement to T, and so by Procrastinate it is deferred until LF—making the directionality of Tense irrelevant to the surface (final) position of verbs. The contrasting structures at Spell Out are illustrated below:

(41)   a.   Imperfective   b.   Perfective

```
            AspP                                           AspP
           /    \                                         /    \
        VP      Asp [IMPF]                             VP      Asp [PERF]
       /  \                                           /  \        |
    ADJ    V                                       ADJ    V      do_i
    broken do                                      broken t_i
```
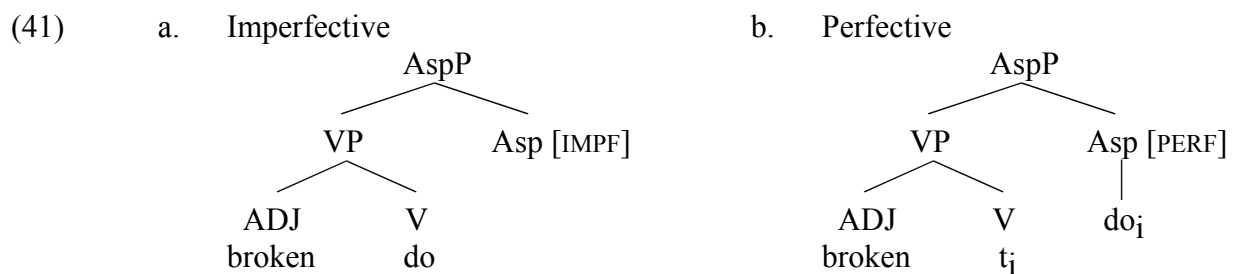
When Spell Out representations are mapped to PF representations, compound verbs will therefore derive different prosodic structures, depending on their perfectivity. At Spell Out, the two elements of an imperfective compound verb form a single phonological word ω. In a perfective compound verb, however, the two elements are divided by a maximal projection (VP) after movement of the auxiliary verb to Asp, and so they form separate phonological words. After the syntactic structure is mapped to prosodic structure and eliminated (syntactic features being uninterpretable at this level), the following PF representations emerge:

(42)   a.   Imperfective   b.   Perfective

              [$_\omega$ broken do]   [$_\omega$ broken] [$_\omega$ do]
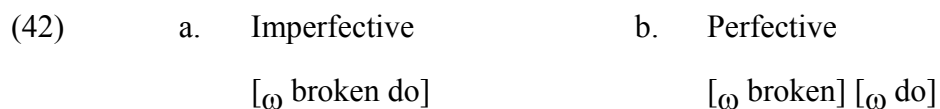
Three types of evidence from the behavior of compound verbs support this analysis: a morphophonological process of merger and its consequences for clitic placement (treated together in section 2.2), and split agreement (section 2.3). Details of the proposed analysis are given in

section 2.4. Section 2.5 takes stock of the range of syntactic projections, setting the stage for a detailed examination of second-position clitic placement in the remaining chapters.

2.2. Merger and clitic placement

Insight into the nature of the relation between the two parts of the compound verb is offered by a morphophonological process that permits the parts of the compound to be merged into a single word. In the imperfective forms of compound verbs, when the noun or adjective ends with a consonant, the initial *k* of the following auxiliary is dropped, and the rest of the auxiliary is added to the noun or adjective to form a single word (Tegey and Robson 1996: 109). This is illustrated below for the compound verb 'to injure', formed from the adjective *dzhóbel* 'injured' and the 1sg forms of the transitive auxiliary *kaw-* 'to make, to do' for the four basic alternations of tense and aspect. All of the sentences in this section are from Yusufzai Pashto, and contain the second-position 2sg clitic *dee*, which is underlined throughout; its position offers clues regarding the structure of the compound verbs:

(43)   Present imperfective

dzhobl-áw-em   <u>dee</u>

injure-TRANS(PRES IMPF)-1SG   *2SG*

'I am injuring you'

(44)   Past imperfective

dzhobl-awél-em   <u>dee</u>

injure-TRANS(PAST IMPF)-1SG   *2SG*

'you were injuring me'

(45)   Present perfective[17]

    a.   dzhóbel   k-em   <u>dee</u>

          injured   *do(PRES PERF TRANS)-1SG*   *2SG*

          'I injure you'

    b.   dzhóbel   <u>dee</u>   k-em

          injured   *2SG*   *do(PRES PERF TRANS)-1SG*

          'I injure you'

(46)   Past perfective

    a.   dzhóbel   krr-em   <u>dee</u>

          injured   *do(PAST PERF TRANS)-1SG*   *2SG*

          'you injured me'

    b.   dzhóbel   <u>dee</u>   krr-em

          injured   *2SG*   *do(PAST PERF TRANS)-1SG*

          'you injured me'

In the imperfective sentences of (43–44), the initial *k* of the auxiliary is omitted, and the compound verb forms a single word, which must be followed by the second-position 2sg clitic *dee*. (The clitic cannot precede the verb unless there is a sentence-initial constituent to host the clitic; see the following two chapters for ample illustration of clitic placement.) In the perfective sentences of (45–46), however, the initial *k* of the auxiliary is retained; the two parts of the verb

---

[17] There is no good explanation at present for why the clitic may follow either the first or second word of the perfective forms, but the crucial point is illustrated by the (b) variants: simply that perfective compounds may be divided by clitics, whereas imperfective compounds may never be divided. This point will be more strikingly illustrated in the next set of sentences.

remain separate, and the 2sg clitic *dee* may either follow the complex verb or appear between its two parts.[18]

Initially, it might appear that the possibility of dropping the initial consonant of the auxiliary and fusing the compound verb into a single verb is nothing more than an artifact of the particular phonological forms within the auxiliary paradigm. In particular, fusion could not apply to the perfective auxiliaries, since their stems generally comprise a single consonant (see their forms in (38–39) above), and so deletion of those onset positions would render the perfective forms identical for all the auxiliaries. In contrast, the imperfective auxiliary stems retain distinct vowels and consonants even after their initial *k* is deleted. Nevertheless, aspect itself may be seen more clearly to play a role in morphological fusion when the placement of second-position clitics is considered. As was shown in (45–46), the second-position 2sg clitic *dee* may appear between the constituents of a compound verb when they form separate words. If this behavior were solely the result of morphology or phonology determining when the initial *k* of the auxiliary must be retained, it would be expected that in compound verbs derived from an adjective ending in a vowel (which never permit the initial *k* of the auxiliary to be deleted), the clitic would similarly be able to divide the constituents of the compound verb. As the following sentences show, however, this is not the case. The clitic may divide the constituents of a compound verb only when it is in perfective aspect:

(47)   Present imperfective

    a.   khaaysta   kaw-em   <u>dee</u>

        beautiful   *do(PRES IMPF TRANS)-1SG*   *2SG*

        'I am making you beautiful'

---

[18] Tegey's (1977: 98) variety of Pashto requires that the initial constituent of the compound verb host the clitic when it is stressed. Note also in these sentences that ergative interpretations occur specifically in past tense, rather than in perfective aspect, as occurs in Hindi (Mahajan 1990: 72–73).

b.   *khaaysta   <u>dee</u>   kaw-em

        beautiful   *2SG*   *do(PRES IMPF TRANS)-1SG*

        'I am making you beautiful'

(48)   Past imperfective

a.   khaaysta   kawél-em   <u>dee</u>

        beautiful   *do(PAST IMPF TRANS)-1SG*   *2SG*

        'you were making me beautiful'

b.   *khaaysta   <u>dee</u>   kawél-em

        beautiful   *2SG*   *do(PAST IMPF TRANS)-1SG*

        'you were making me beautiful'

(49)   Present perfective

a.   khaaysta   k-em   <u>dee</u>

        beautiful   *do(PRES PERF TRANS)-1SG*   *2SG*

        'I make you beautiful'

b.   khaaysta   <u>dee</u>   k-em

        beautiful   *2SG*   *do(PRES PERF TRANS)-1SG*

        'I make you beautiful'

(50)   Past perfective

a.   khaaysta   krr-em   <u>dee</u>

        beautiful   *do(PAST PERF TRANS)-1SG*   *2SG*

        'you made me beautiful'

b.   khaaysta   <u>dee</u>   krr-em

      beautiful   *2SG*   *do(PAST PERF TRANS)-1SG*

      'you made me beautiful'

Imperfective compound verbs thus behave as a single unit with respect to clitic placement, irrespective of whether the initial *k* of their auxiliary is deleted. Aspect itself thus plays a role in the formation of compound verbs, with imperfective (but not perfective) compound verbs being impenetrable to second-position clitics.[19]

2.3. Split agreement

Compound verbs show that agreement is yet more complicated than suggested by the introductory remarks in chapter 1, since the two parts of the compound verb may agree with different constituents in the same sentence. Both parts of the compound verb agree with the object in past perfective transitive sentences, as might be expected given the pattern of ergativity that was illustrated in (34) above with a simple verb. The sentences in this section are from Yusufzai:[20]

(51)   Past perfective: object agreement

      a.   sangin   kerkey   maat-a   krr-a

            Sangin(MASC)   *window(FEM SG)*   *broken-FEM SG*   *do(PAST PERF)-FEM3SG*

            'Sangin broke the window'

---

[19] Tegey (1977: 98–99) notes the perfective/imperfective distinction, imputing it solely to stress, but the adjectival portion of all of these verbs bears final stress.

[20] Thanks to Jan Mohammad for suggesting these paradigms.

b. sangin   kerkey   maat-ee   krr-i

Sangin(MASC)   *windows(FEM PL)* *broken-FEM PL* *do(PAST PERF)-FEM3PL*

'Sangin broke the windows'

c. sangin   war   maat   krr-o

Sangin(MASC)   *door(MASC SG)*   *broken(MASC SG)* *do(PAST PERF)-MASC3SG*

'Sangin broke the door'

d. sangin   warúna   maat   krr-el

Sangin(MASC)   *doors(MASC PL)* *broken(MASC PL)* *do(PAST PERF)-MASC3PL*

'Sangin broke the doors'

At this point, the two parts of the compound verb could be regarded as a single lexical item that

agrees with the object. Such a conclusion would also be supported by the agreement pattern in

the present and past imperfective examples below, in which the two parts of the compound verb

form a single word, and the adjectival portion is uninflected:

(52)   Present imperfective: subject agreement

a. sangin   kerkey   maat-aw-i

Sangin(MASC)   *window(FEM)*   *broken-TRANS(PRES IMPF)-3SG*

'Sangin is breaking the window(s)'

b. sangin   war   maat-aw-i

Sangin(MASC)   *door(MASC SG)*   *broken-TRANS(PRES IMPF)-3SG*

'Sangin is breaking the door'

c. sangin　　　　　warúna　　　　　maat-aw-i

Sangin(MASC)　　doors(MASC PL)　　broken-TRANS(PRES IMPF)-3SG

'Sangin is breaking the doors'

d. táaso　　　　　warúna　　　　　maat-aw-ey

2PL　　　　　doors(MASC PL)　　broken-TRANS(PRES IMPF)-2PL

'you (PL) are breaking the doors'

(53)　Past imperfective: object agreement

a. Sangin　　　　　kerkey　　　　　maat-aw-éla

Sangin(MASC)　　window(FEM SG)　　broken-TRANS-FEM3SG(PAST IMPF)

'Sangin was breaking the window'

b. Sangin　　　　　kerkey　　　　　maat-aw-éli

Sangin(MASC)　　windows(FEM PL)　　broken-TRANS-FEM3PL(PAST IMPF)

'Sangin was breaking the windows'

c. Sangin　　　　　war　　　　　maat-aw-u

Sangin(MASC)　　door(MASC SG)　　broken-TRANS-MASC3SG(PAST IMPF)

'Sangin was breaking the door'

d. Sangin　　　　　warúna　　　　　maat-aw-el

Sangin(MASC)　　doors(MASC PL)　　broken-TRANS-MASC3PL(PAST IMPF)

'Sangin was breaking the doors'

As explained in the previous section, since the adjectival stem of the compound verb ends in a consonant (*maat* 'broken'), the initial *k* of the transitive auxiliary *kaw-* is omitted, and the final

consonant of the adjective forms the onset of the following syllable. The invariant form of the adjective shows that it does not agree with the object in either present or past imperfective, although the entire (derived) verb agrees with the subject in present tense (52), and with the object in past tense (53)—the familiar pattern of tense-split ergativity.

Evidence for disassociating subject and object agreement in a single sentence, however, comes from perfective aspect in non-past tense sentences, in which the adjectival portion of the compound verb agrees with the object, while the perfective auxiliary agrees with the subject. Present tense sentences are given in (54). The paradigm in (55) gives future tense sentences, which are created from a past perfective auxiliary and the second-position future clitic *ba*. These sentences show that the auxiliary verb behaves independently, and according to the usual, ergative pattern, agreeing with the object in past tense, and with the subject in non-past tense.

(54)   Present perfective: split agreement

    a.   táaso   kerkey   maat-a   key

        2PL   *window(FEM SG)*   *broken-FEM SG*   *do(PRES PERF 2PL)*

        'you (PL) break the window'

    b.   táaso   kerkey   maat-ee   key

        2PL   *windows(FEM PL)*   *broken-FEM PL*   *do(PRES PERF 2PL)*

        'you (PL) break the windows'

    c.   táaso   war   maat   key

        2PL   *door(MASC SG)*   *broken(MASC SG)*   *do(PRES PERF 2PL)*

        'you (PL) break the door'

d. táaso   warúna   maat   key

2PL   *doors(MASC PL)*   *broken(MASC PL)*   *do(PRES PERF 2PL)*

'you (PL) break the doors'

(55)   Future: split agreement

a.   sangin   <u>ba</u>   kerkey   maat-a   krr-i

Sangin   *FUT*   *window(FEM SG)* *broken-FEM SG*   *do(PAST PERF)-MASC3SG*

'Sangin (MASC) will break the window'

b.   sangin   <u>ba</u>   kerkey   maat-ee   krr-i

Sangin   *FUT*   *windows(FEM PL)* *broken-FEM PL*   *do(PAST PERF)-MASC3SG*

'Sangin (MASC) will break the windows'

c.   sangin   <u>ba</u>   war   maat   krr-i

Sangin   *FUT*   *door(MASC SG)*   *broken(MASC SG)*   *do(PAST PERF)-MASC3SG*

'Sangin (MASC) will break the door'

d.   sangin   <u>ba</u>   warúna   maat   krr-i

Sangin   *FUT*   *doors(MASC PL)*   *broken(MASC PL)*   *do(PAST PERF)-MASC3SG*

'Sangin (MASC) will break the doors'

e.   táaso   <u>ba</u>   warúna   maat   krr-ey

2PL   *FUT*   *doors(MASC PL)*   *broken(MASC PL)*   *do(PAST PERF)-2PL*

'you (PL) will break the doors'

This split agreement pattern also appears in the imperative mood of compound verbs, which are similarly formed from both present and past perfective auxiliaries. When the imperative has a

single addressee, the singular imperative *-a* appears on the auxiliary, regardless of the gender or number of the object, while the adjectival portion of the compound verb varies with the gender and number of the intended object:[21]

(56)   Imperative: split agreement

    a.   dzhorr   k-a

          built(MASC SG)   *do(PRES PERF)-2SG(IMP)*

          'build it(MASC SG)!'

    b.   dzhórra   k-a

          built(FEM SG)   *do(PRES PERF)-2SG(IMP)*

          'build it(FEM SG)!'

    c.   dzhorr   k-a

          built(MASC PL)   *do(PRES PERF)-2SG(IMP)*

          'build it(MASC PL)!'

    d.   dzhórri   k-a

          built(FEM PL)   *do(PRES PERF)-2SG(IMP)*

          'build it(FEM PL)!'

When the imperative has a plural addressee, the auxiliary bears the 2pl suffix *-ey*, while the adjective continues to agree with the intended object:

---

[21] This paradigm is suggested by Tegey and Robson (1996: 131–132), but the actual forms are from Yusufzai. The use of a past- rather than a present-tense verb for a plural addressee appears to be a quirk of Yusufzai, as the reference above states that the present perfective stem of the verb takes both singular (*-a*) and plural (*-ey*) imperative suffixes.

(57)   Imperative: split agreement

   a.   dzhorr   krr-ey

         built(MASC SG)   do(PAST PERF)-2PL

         'build it(MASC SG)!'

   b.   dzhórra   krr-ey

         built(FEM SG)   do(PAST PERF)-2PL

         'build it(FEM SG)!'

   c.   dzhorr   krr-ey

         built(MASC PL)   do(PAST PERF)-2PL

         'build it(MASC PL)!'

   d.   dzhórri   krr-ey

         built(FEM PL)   do(PAST PERF)-2PL

         'build it(FEM PL)!'

Unlike agreement in indicative CVs, the tense of the auxiliary verb is irrelevant to agreement.

An outline of the somewhat complicated constituency of compound verbs has emerged. The patterns illustrated above are summarized in the following table. Note that the NV (adjectival or nominal) element of the compound verb is either uninflected, or agrees with the object; unlike the LV (auxiliary), adjectives never show subject-agreement, regardless or tense, aspect, or mood. Auxiliaries, on the other hand, must always agree with either the subject or object.

(58)		Agreement in compound V	Mood	Aspect	Tense	Example
	a.	Adj + Aux$_{SUBJECT}$	Indic	Imperfective	Present	(52)
	b.	Adj + Aux$_{OBJECT}$	Indic	Imperfective	Past	(53)
	c.	Adj$_{OBJECT}$ Aux$_{OBJECT}$	Indic	Perfective	Past	(51)
	d.	Adj$_{OBJECT}$ Aux$_{SUBJECT}$	Indic	Perfective	Present	(54)
			Indic	Perfective	Future	(55)
			Imp	Perfective	Present	(56)
			Imp	Perfective	Past	(57)

As suggested in the previous section, imperfective compound verbs (58a–b) behave as a single lexical item, which is why their adjectival portion is invariably uninflected. The choice of subject vs. object agreement on the auxiliary is the usual one determined by tense: subject agreement in present tense, and object agreement in past tense. In past perfective compounds (58c), both the adjective and auxiliary agree with the object, as is to be expected in past tense, which always shows ergativity. The fact that the adjective in perfectives is inflected at all, though, distinguishes it from its counterpart in imperfectives, and shows that the two parts of the compound verb behave differently. Although both parts here agree with the object (since the verb is past tense), evidence for their separate status comes from non-past perfectives and imperatives (58d), since in those constructions, the two parts of the compound verb agree with different constituents of the sentence.

It was mentioned in chapter 1 that Pashto has seemed unlike Hindi/Urdu in defining its ergative split on tense, rather than aspect. Agreement in compound verbs shows that Pashto nevertheless retains an element of aspect-driven ergativity, since adjectival object agreement is indeed defined on perfective aspect (as in Hindi/Urdu), rather than on tense. It is only the auxiliary component of the compound verb that exhibits ergativity in past tense. Agreement and case-marking thus do not exhibit a single pattern of ergativity in Pashto, as the behavior of simple

verbs would suggest. Explaining this disassociation of agreement is the purpose of the next section.

2.4. Explaining the asymmetries

2.4.1. Argument structure

Before discussing the defining split between perfective and imperfective compound verbs, some remarks about the argument structure of predicates are in order. The properties of argument structure will also be crucial in understanding the licensing and interpretation of second-position clitics, to be discussed in the next chapter.

Pashto's nearest modern sister, Persian (Farsi), is similar to Pashto in having compound verbs, although Persian has a considerably larger array of verbs that may appear as the second element of the compound; in Pashto, only auxiliary verbs may form the second member of the compound.[22] Karimi's (1997a,b) recent analysis of compound verbs (CVs) suggests that they contain two parts: the first, non-verbal (NV) element, and the second, verbal element (the light verb: LV), in the following configuration:

(59)
```
              CV
            /    \
         XP        LV
         |
         X
```
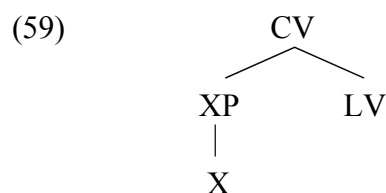
At some point in the derivation—overtly or covertly—the head of the NV (represented here as X, as its category is open) adjoins to the LV:

---

[22] Babrakzai (1999: 140–141) identifies a couple of exceptions.

(60)

```
              CV
            /    \
          XP      LV
          |      /   \
          ti    Xi    LV
```
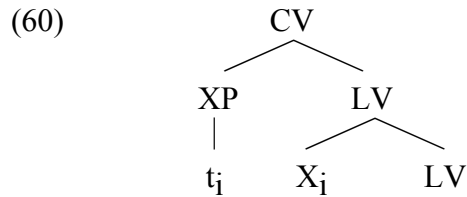
While this structure expresses the observation that the two parts of a CV form a single constituent, it is not able to accommodate compound verbs in Pashto, since—as was shown in section 2.3—the two parts of a compound verb may behave either as two units or one; in the former case (i.e., perfective aspect), the constituents of the CV will need to be kept distinct for purposes of agreement.

Another approach that permits the needed distinctions is to employ a lexical argument structure of the kind suggested by Hale and Keyser (1993), which explicitly encodes transitivity relations. Since intransitive CVs are formed with the intransitive auxiliary *keeg-* 'to become', and transitive CVs with the auxiliary *kaw-* 'to make, to do', such verbs may be derived by successive adjunction. For example, when combining an adjective with the intransitivizer *keeg-*, the adjective projects an internal argument position as specifier of VP; the argument that is merged into this specifier position, DP_1, ultimately will become the direct object in the transitive sentence that is being constructed:[23]

(61)

```
              VP
            /    \
          DP1     V'
                 /   \
               ADJ    V
                   keeg- INTR
```

---
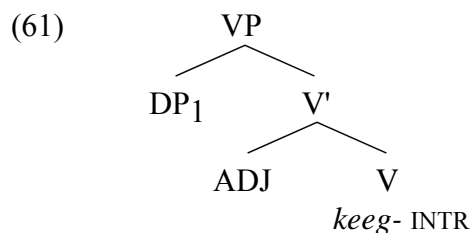
[23] For convenience, labels for the verbal categories are given as VP, V', and V, etc., although they are more precisely to be regarded as projections of the abstract (in)transitivity features that form the heads of those categories.

The verb thus derived may be made transitive by having the structure in (61) become the complement of the features associated with transitive *kaw-*, which occupy *v*. As shown in (62), the transitivizer projects an external argument DP$_2$ (the subject) in Spec/*v*P.

(62)

```
                    vP
            ┌───────┴───────┐
          DP₂              v′
                   ┌────────┴────────┐
                  VP                 v
             ┌─────┴─────┐        kaw- TRANS
           DP₁          V′
                   ┌─────┴─────┐
                  ADJ          V
                          keeg- INTR
```
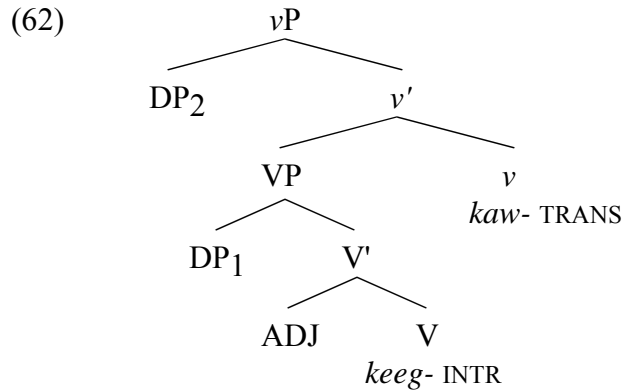
Only the intransitive and transitive features associated with *keeg-* and *kaw-* are combined in this manner—not their phonological form; their co-occurrence is spelled out as *kaw-* under V, along with features of tense, aspect, and agreement.[24] The structure in (62) would itself be selected by the category projected by aspectual features—alluded to in (40)—and so on, finally yielding a clause structure articulated as follows:

---

[24] Similar constructions in other languages share this underlying structure, but select their own arbitrary Spell Out forms, in accord with the observation that language variation resides in the lexicon rather than in the syntax (Chomsky 1995: 169–170).

(63)

```
                    CP
              ┌──────┴──────┐
          C [MOOD]          TP
                      ┌──────┴──────┐
                      T            AspP
               [PAST/NON-PAST]  ┌───┴────┐
                               vP    Asp [PERF/IMPF]
                           ┌────┴────┐
                          DP₂        v'
                                ┌────┴────┐
                               VP      v [kaw- TRANS]
                           ┌────┴────┐
                          DP₁        V'
                                ┌────┴────┐
                               ADJ    V [keeg- INTR]
```
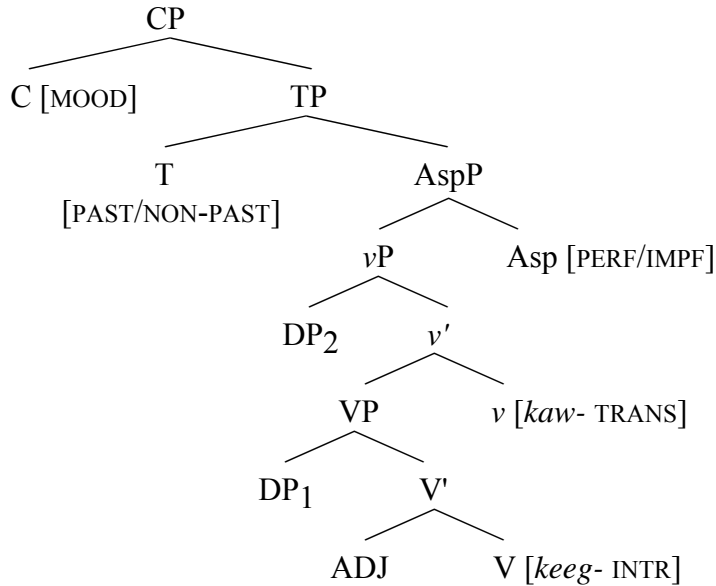
The highest category C has already been shown in the discussion of (13–17) to be head-initial, unlike the lexical categories. Tense is shown as head-initial, in keeping with several other functional projections—though it should again be noted that the directionality of Tense does not affect the surface word order of the verb, which enters the derivation fully inflected from the lexicon and remains head-final at Spell Out. There is no need for overt movement to T, and so by Procrastinate it is deferred until LF—making the directionality of Tense irrelevant to the surface word order of verbs. The clause also contains agreement projections, but these will not be introduced until the next chapter.

2.4.2. Perfective as a strong feature

Aspect has been shown to be the feature that most directly affects the derivation of compound verbs: when imperfective, the adjectival element of the compound verb is uninflected; when perfective, the adjectival element agrees with the object. The auxiliary element of the compound verb agrees behaves independently (as if it were a simple, main verb), agreeing with the object in past tense, and the subject in non-past tense.

Aspect projects between TP and VP, as was illustrated above in (40) and (63), and is headed either by perfective or imperfective features. For simple verbs, perfective features are also spelled out as the morpheme *we*. Inflected verbs enter the numeration fully inflected from the lexicon. Because perfective aspect is a strong feature, it triggers overt movement of the verb to Asp, where the perfective feature is erased. In the case of compound verbs, only the second element, the auxiliary, moves to Asp, due to the Head Movement Constraint.[25] Imperfective aspect, a weak feature, also triggers movement, but it is deferred until after Spell Out (by Procrastinate). The differing representations of compound verbs at Spell Out were illustrated above in (41), which is repeated here as (64); for brevity, the intervening *v* is omitted from each structure.

(64)　　a.　Imperfective　　　　　　　　　　b.　Perfective

```
            AspP                                      AspP
           /    \                                    /    \
         VP      Asp [IMPF]                        VP      Asp [PERF]
        /  \                                      /  \        |
     ADJ    V                                  ADJ    V      do_i
    broken  do                                broken  t_i
```
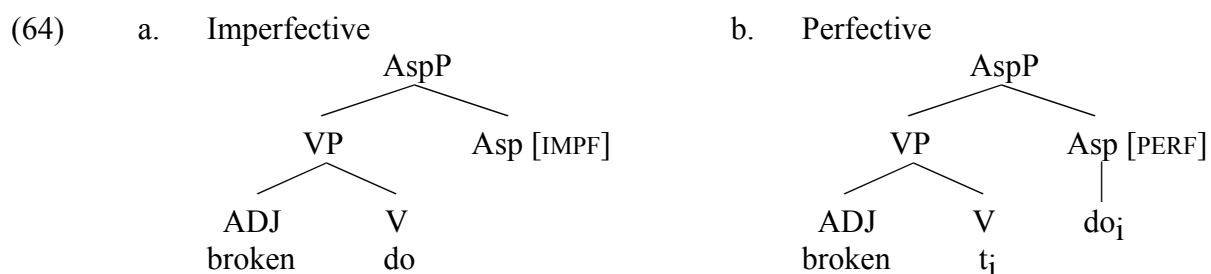
These structures explain several otherwise puzzling facts. First, as was discussed in section 2.2, imperfective compound verbs are subject to a phonological rule that deletes the initial consonant of the auxiliary component of the compound verb, merging the two components into a single word. Assuming that this rule does not apply across a maximal projection (see Selkirk 1984, 1986 on the syntax-phonology mapping), the structures in (64) explain why this rule does not apply to perfective compound verbs: they have a maximal projection (VP) intervening between the two parts of the compound verb before Spell Out, which is the representation that feeds the level of PF (Phonological Form); recall their differing representations above in (42). For a similar reason, these structures also explain why clitics may divide the two parts of a perfective

---

[25] 'An X^0 may only move into the Y^0 which properly governs it' (Travis 1984: 131).

compound verb: because at Spell Out they do not appear together within VP, and therefore derive distinct phonological words.

The split-agreement pattern of compound verbs is also related to aspect, since it is specifically perfective aspect (which is morphologically represented only on the LV) that triggers agreement between the direct object and the NV element of the compound. This pattern bears a superficial resemblance to past participle agreement in Standard French, in which the participle agrees with pre-verbal objects (specifically, accusative clitics and *wh*-trace), but not with post-verbal objects, as illustrated below:

(65)   a.   Paul les a repeint**es**

'Paul has repainted them'

b.   les chaises que Paul a repeint**es**

'the chairs that Paul has repainted'

c.   Paul a repeint(***es**) les chaises

'Paul has repainted the chairs' (Kayne 1989a: 85)

This correlation between NP position and agreement in French has been taken to suggest that agreement is triggered by the direct object having moved to, or through, AgrO (Kayne 1989a, Van Gelderen 1997, Déprez 1998). There is considerable variation in this pattern both within and without Romance, for which see the above references. In Pashto, of course, the verb rigidly follows the direct object, and so there is no correlation between movement of the object and agreement, as there is in Standard French. Furthermore, the direct object in Pashto is always assumed to move into the functional projections of the clause at Spell Out (in order to precede the aspect and negation projections, and erase person/number features in AgrO), and so movement (or not) of the direct object as an explanation of split agreement is not tenable. Instead, given the

analysis sketched above in which perfective auxiliaries move to Asp, split agreement should instead be correlated with movement of this head. This analysis entails that there is an agreement feature of the direct object that is erased within VP. The structure of a compound verb is as follows, in which DP$_1$ is the position of the direct object, ADJ is the position of the NV element of the compound verb, and V is the original position of the LV (before it raises to *v* and the higher heads).

(66)

```
                        AspP
                     ／      ＼
                  vP          Asp [PERF/IMPF]
               ／    ＼
           DP₂          v′
                     ／    ＼
                  VP          v [kaw- TRANS]
               ／    ＼
           DP₁          V′
                     ／    ＼
                  ADJ          V [keeg- INTR]
```
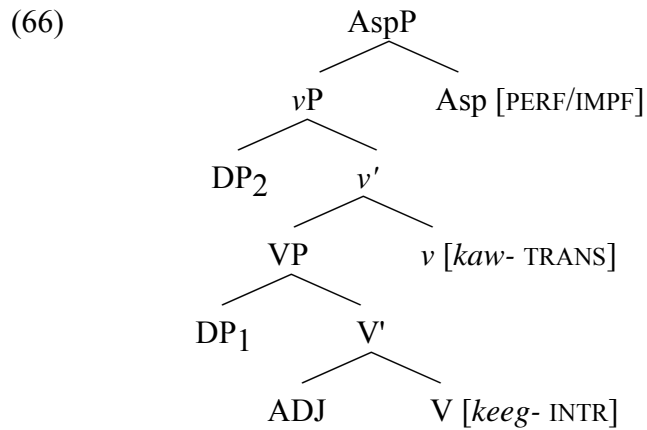
In imperfective aspect, the direct object, DP$_1$, erases its agreement feature with the LV under V, and so ADJ remains bare. In perfective aspect, however, the LV must raise to Asp in order to erase the strong perfective feature, and so it may not erase agreement with the direct object. In this case, the only element in a local relation with DP$_1$ is the NV (ADJ), and so the NV must bear the agreement feature, which is morphologically realized on the NV. In both perfective and imperfective aspect, the LV behaves independently as if it were a main verb, showing overt agreement with the object in past tense, and with the subject in non-past tense.

Something additional needs to be said about split agreement in imperatives. Recall from the summary in (58) that although imperative compound verbs may be formed with a past tense auxiliary, the auxiliary does not exhibit ergativity, i.e., the auxiliary agrees with the subject, rather than with the object, as occurs in indicative mood (58b–c). A possible approach to this splitting of agreement in imperatives might be to assume that an imperative operator in C$_0$ forces verb

movement (Han 1998); on its way to C, the auxiliary would pass through Tense, agreeing with the 2sg *pro* subject in its specifier. However, because CP is head-initial, as was shown in the discussion of (13–17) in the previous chapter, this derivation would not produce the desired verb-final word order. For this reason, it will be supposed instead that the suffixes on imperative verbs do not represent canonical agreement morphology. Recall the paradigm of verbal agreement suffixes in (30) in chapter 1. Although plural imperatives (57) bear the regular 2pl agreement suffix *-ey*, singular imperatives (56) do not bear the ordinary 2sg agreement suffix *-ee*, but rather bear the suffix *-a*, which is restricted to imperative mood. Because imperatives are not comparable to main clauses (neither in their tense determining agreement, nor in the suffixes they bear), and because imperatives have a special discourse function with second-person addressees (as in all languages), it will be assumed that imperative verbs do not participate in ordinary processes of agreement. In other words, the suffixes that appear on imperative verbs are not subject agreement, but imperative morphology.

2.5. Split headedness

Now that aspect has been examined closely, we are able to flesh out the remaining functional structure of the clause. It may be noted that because the perfective morpheme *we* precedes the negative morpheme *ne*, which in turn always precedes the verb—as exemplified below; see also (31) above—NegP occupies a position between AspP and VP:

(67)   ahmad   <u>ba</u>   [AspP we   [NegP ne   [VP gaddeeg-i]]]
          Ahmad   *FUT*   *PERF*   *NEG*   *dance-3SG*
          'Ahmad will not dance' (Tegey and Robson 1996: 128)

In a transitive sentence with full NPs, the object precedes AspP:

(68)   asad   <u>ba</u>   lik   [$_{AspP}$ we   [$_{NegP}$ ne   [$_{VP}$ leeg-i]]]

          Asad   *FUT*   *letter*   *PERF*   *NEG*   *send-3SG*

          'Asad will not send the letter' (Tegey and Robson 1996: 128)

This is taken to be evidence of scrambling of the object out of VP to the specifier of an intervening agreement projection—also head-initial—which will be discussed in the next chapter. In the meantime, the clause now takes the form in (69). As has been mentioned at various points, lexical categories are rigidly head-final, while functional categories are not.

(69)
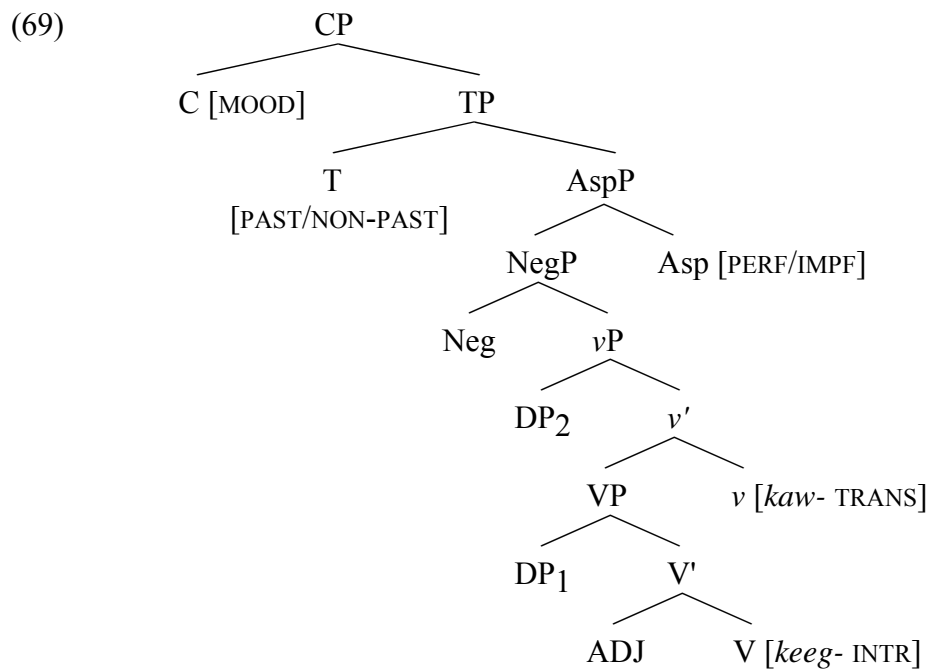
There is contradictory evidence for the headedness of AspP. Previous sections suggested that the perfective auxiliary of compound verbs moves to Asp to erase the strong perfective feature there. Because the auxiliary remains final at Spell Out, AspP was assumed to be head-final, as was schematized in (64). But considering that the perfective morpheme *we* precedes negation, as was shown in (67–68), might not AspP instead be regarded as head-initial, in an attempt to bring it in line with other head-initial functional categories (CP, DP, AgrP, NegP)?

A solution to this contradiction is to continue regarding AspP as head-final, and to specify the perfective morpheme *we* as a proclitic in its lexical entry:

(70)   Interpretation: perfective

         Category: Asp

         Phonetic Form: */we=/*

The categorical information probably need not be explicitly included in the lexical entry, being derivable from the interpretation, which is aspectual. At PF (or in an autonomous morphological component), this morpheme's status as a proclitic results in its being inverted, so that it appears on the other side of NegP and VP, as required. Note that the category Asp does not invert—it remains final throughout the derivation—but that only the order of morphemes is altered. This is a purely post-syntactic operation, similar to the prosodic inversion of second-position clitics to be discussed in chapter 4.

Corroborating evidence for this analysis—as well as for the analysis of perfective aspect itself as a strong feature—comes from a class of verbs that comprise a bound stem and an historic prefix. One such verb is *kee-naastel* 'to sit down', which comprises the semantically opaque, historical prefix *kee-* and the stem *naastel*; another is *pree-khodel* 'to leave/abandon', which comprises the historical prefix *pree-* and the stem *khodel*. Verbs in this category do not take the otherwise productive *we* morpheme to form the perfective, but rather derive their perfective form by shifting stress to the initial syllable (i.e., onto the prefix), as illustrated below:

(71)   a.   Imperfective

                    shaageerdaan   kee-naast-él

                    students   PREFIX-sit down-MASC3PL

                    'The students were sitting down'

b.   Perfective

    shaageerdaan   kée-naast-el

    students   PREFIX-sit down-MASC3PL

    'The students sat down' (Babrakzai 1999: 53)

If the negative morpheme *ne* is added to imperfective (71a), it attracts the stress (as happens also with simple verbs):

(72)   Imperfective

    shaageerdaan   né-kee-naast-el

    students   NEG-PREFIX-sit down-MASC3PL

    'The students were not sitting down' (Babrakzai 1999: 53)

A striking contrast occurs, however, when the negative *ne* morpheme is added to perfective (71b). The negative morpheme attracts the stress, but the historical 'prefix' *kee-* now precedes negation:

(73)   Perfective

    shaageerdaan   kee-né-naast-el

    students   PREFIX-NEG-sit down-MASC3PL

    'The students did not sit down' (Babrakzai 1999: 53)

It was just shown in (67–68) that AspP is higher than NegP, and a contrast like (72–73) offers further evidence not only for the higher position of AspP, but for perfective being a strong feature under its head, compelling movement there—crucially, the minimal movement that will erase the strong perfective feature. In the case of historically complex verbs like *kee-naastel* 'to sit down', the morphological structure of these verbs is loose enough that the apparent prefix (its synchronic

behavior is better regarded as proclitic) may separate from the verb stem, moving to Asp to erase the strong perfective feature.[26]   In other words, the historic prefix in (73) appears in the same position as the regular, perfective *we* morpheme in (67–68). Their complementary distribution is explained, because they erase a strong perfective feature in the same position: the Asp head. In the case of regular compound verbs (formed from a noun or an adjective plus an auxiliary), it is the auxiliary (light verb) element that moves to Asp, due to the Head Movement Constraint. At Spell Out, then, the structure of the relevant portion of (73) is as follows:

(74)

```
                        AspP
                _____/    _____
             NegP                     Asp
           __/    \__                  |
         Neg       VP               kee-ᵢ
          |      __/\__
          ne   tᵢ naastel 'sit' (MASC3PL)
```
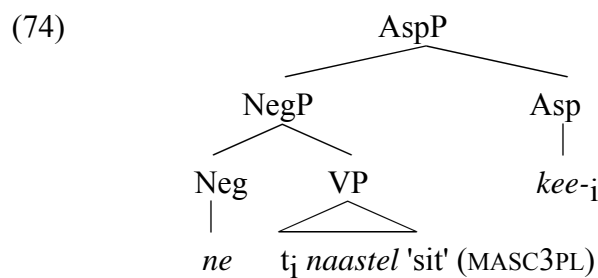
At PF, the syntactic information is stripped away in the process of deriving the prosodic structure as fully as possible, with each syntactic head forming a phonological word. Because negative *ne* is itself a proclitic, it forms a word with its host, the verb stem, and together they form a single stress-bearing unit:

(75)   [$_\omega$ [$_\omega$ ne] naastel] kee-

Because the *kee* is proclitic, it lacks a phonological host in (75), and so it undergoes minimal (leftward) prosodic inversion (Halpern 1995), a last resort operation in which the proclitic may be licensed at PF:

---

[26] Another possibility is that separable prefixes are so-called preverbs, which exist in many languages and show a variety of behaviors, including separability from the stems with which they are associated (Ackerman and Webelhuth 1998).

(76)   [kee- [$_\omega$ [$_\omega$ ne] naastel]]

In this structure, the *kee-* morpheme has both of its requirements satisfied: at Spell Out, it has erased the strong perfective feature under the Asp node, and at PF it finds a suitable phonological host. The different surface orders of imperfective (72) versus perfective (73) are thus explained.

The same point may be demonstrated by the placement of second-position clitics with respect to these verbs. As shown below for the historically complex verb *pree-khodel* 'to leave/abandon', the second-position clitic 2sg *dee* may divide the verb stem when it is perfective, as it is in the following two sentences:

(77)   a.   taa   kaalina   prée-khod-a

                  PN2SG   *rug(FEM)*   *PREFIX-leave-FEM3SG*

                  'Did you leave the rug?'

            b.   prée   <u>dee</u>   khod-a

                  PREFIX   *2SG*   *leave-FEM3SG*

                  'Did you leave it?' (Babrakzai 1999: 54)

Because both sentences are perfective, the prefix *pree* has moved to Asp to erase the strong perfective feature. In (77a), the 2sg strong pronoun subject *taa* is in Spec/TP, while the object has scrambled outside of VP. It is only in (77b), though, that movement of the prefix may be seen clearly; here, the object *kaalina* 'rug' from (77a) has been omitted, as it may be identified by the object agreement suffix on the verb. The strong 2sg pronoun *taa* from (77a) now takes the form of a clitic; because it requires a phonological host to its left, the clitic minimally inverts, to the right of the historic prefix, which has independently separated from the verb stem in order to erase the strong perfective feature.

By assuming a head-final AspP, and assigning historical prefixes lexical entries of the kind in (70), i.e., treating them as proclitic, the headedness of the functional projection AspP makes it appear more like the lexical categories NP and VP. It is not clear why AspP should be one of the few functional categories that appears to be head-final. One possibility is that Asp is more 'lexical' than the other functional categories, in that it may be overtly targeted by a light verb (as discussed in previous sections); none of the other functional categories behave in this way.

The apparent mixed nature of AspP is similar to that of PP—alluded to much earlier—but the latter is also amenable to a uniform treatment. Because Pashto has prepositions, postpositions, and ambipositions (also called 'circumpositions'), it initially seems difficult to establish a single head-directionality for this category. Examples are given below (Tegey and Robson 1996: 153–155):

(78)   a.   **Prepositions**

dee   'of (POSS)'

pe   'by means of, with; at (time expressions)'

    b.   **Postposition**

ta   'to (DAT)'

    c.   **Ambipositions**

pe ... kee	'in, at'	*le ... laandee*	'under'
pe ... pesee	'after'	*le ... tsékha*	'from'
pe ... baandee	'on, to'	*ter ... (a) poree*	'up to'
le ... sara	'with'	*ter ... láandee*	'under'
le ... na	'from'		

In ambipositions, the first element (more than the second) tends to be optionally omitted—as illustrated below for *pe ... kee* 'at' and *le ... sara* 'with':

(79)   a.   parun   <u>dee</u>   sona   ghanem

                 yesterday   *2SG*   *how.much*   *wheat(MASC PL)*

                 (pe)   baazaar   **kee**   wáa-khist-el

                 at   *market*   *at*   *PERF-buy-PAST(MASC PL)*

                 'How much wheat did you buy at the market yesterday?' (Laghman)

    b.   laylaa   **(le)**   amaan   **sara**   náasta   da

                 Layla   *with*   *Aman*   *with*   *sitting(FEM SG)*   *be(PRES IMPF FEM3SG)*

                 'Layla is sitting with Aman' (Tegey and Robson 1996: 155–156)

The simplest approach to these constructions is to regard the apparent postpositional element of ambipositions as a lexical category, rather than a functional one, as some of these final elements were indeed lexical in the prehistory of Pashto.[27] If the second element of ambipositions is treated as a lexical category, such phrases receive a simple treatment that respects the general headedness otherwise observed; the ambipositional phrase in (79b), for example, would have the following structure:

---

[27] The morpheme *poree* has a cognate noun in Sanskrit ('the further back of a river'), while *sara* has the cognate noun 'union' in Avestan; the latter is also found in older Pashto literature (Morgenstierne 1927: 58, 69). The second element of at least one ambiposition, *pe ... baandee*, is synchronically an adverb (Penzl 1955: 155; Babrakzai 1999: 42, 46). For discussion of historical antecedents for Pashto ambipositions, see Heston (1987).

(80)

```
              PP
           /      \
          P        NP
          |       /   \
       le 'with' NP     N
                 /\     |
              Amaan   sara 'with'
```
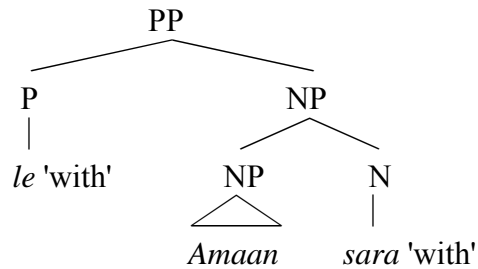
In this structure, the preposition *le* 'with' selects an NP complement, the head of which is the second element of the apparent ambiposition (*sara* 'with'). This lexical head in turn selects its own NP complement, *Aman*, which is the 'object' of the apparent ambipositional phrase. This analysis has the desirable result of maintaining lexical categories as uniformly head-final, as well as treating PP as head-initial, like several other functional categories.[28]

Finally, note that the future morpheme *ba* appears fairly high in the structure of (67–68) above, as it presumably occupies the head of a head-initial ModalP. This suggestion represents a good starting point for the following chapter, concerning clitics, since the future morpheme *ba* is one of the second-position clitics. The remaining chapters will suggest that a group of functional morphemes appears in second position of the clause exactly because they occupy ModalP (and/or similarly high, head-initial functional projections, depending on what clitics happen to co-occur in a cluster), in the same way that second-position clitics in Serbo-Croatian (and the verb in verb-second languages like German) have been assumed to occupy C. The analysis to follow departs from tradition, however, by eliminating the need to stipulate movement of clitics to a functional projection. A number of complications pertaining to second-position clitic placement in Pashto may be explained simply by leaving clitics in their merged positions, and by assuming a clausal structure as in (69): head-final lexical categories and head-initial functional categories. The projections for which there is evidence for head-directionality are listed below.

---

[28] The dative postposition *ta* is an apparent exception. Sec. 3.4.5 below suggests that this postposition is a case-assigner. As verbs also assign case to their left, the dative postposition is plausibly a lexical category (like V), rather than a functional one.

(81)

Category	Direction of head
NP	final
VP	final
AspP	final
PP	initial
NegP	initial
CP	initial
DP	initial
ModalP	initial
AgrP[29]	initial

This table shows that the line dividing the choice of head direction is closely correlated with the one that divides lexical and functional categories.[30]

---

[29] Evidence for head-initial AgrP is presented in ch. 3.

[30] The opposite split has been claimed for American Sign Language, with head-initial lexical categories, and head-final functional categories (Romano 1991). However, as with most interesting syntactic questions, there remains debate about headedness in ASL (Gaurav Mathur, p.c.).

3. Deriving second position

3.1. Introduction

This chapter eliminates the notion of a 'second position' in the Pashto clause in which clitics are hosted, as well as any need for syntactic movement of clitics. The appearance of a group of clitics in second position, and their ordering with respect to each other, is shown to derive from independent principles of syntax and phonology. After adducing evidence that second-position (2P) pronominal clitics are agreement morphemes, rather than arguments, it will be suggested that each clitic heads an agreement projection whose specifier licenses (identifies) an empty pronominal (*pro*), which is the actual argument. This analysis derives the order, position, and interpretation of the clitics without overt syntactic movement of the clitics. A small class of sentences compels movement at PF, but this movement is minimal, and serves the 'last resort' purpose of saving a structure that is well-formed at LF (interpretation), but illicit at PF (phonology). In other words, a derivation that converges at LF will always converge at PF, but the converse is not true.

Pashto's pronominal clitics are typologically unusual in several respects. First, they are divided into two sets, one that appears in second position of the clause, and another that appears nearer the verb. It is far more common for a language to use only one of these positions for its pronominal clitics.[31] Serbo-Croatian, for example, has only second-position clitics, while Romance languages have only verbal clitics. A second unusual property of Pashto is the ordering of pronominal clitics within the cluster, which is strictly determined by grammatical person (rather than by their case or function within the clause), and which appears to be determined by syntax, rather than by a morphological template. Furthermore, the features of the clitics themselves are more impoverished than in other Indo-European languages, distinguishing only

---

[31] Because of this richness of clitics—both in position (second-position vs. verbal) and function (pronominal, modal, and adverbial)—much of the discussion about their status as arguments or agreement will refer, concisely, to 'clitics,' but what is intended is specifically 'second-position pronominal clitics.' Verbal clitics, which carry dative features, are discussed in sec. 3.4.5.

person and number, and yet not even distinguishing these uniquely. As different as Serbo-Croatian and Romance languages are, they nevertheless have in common the ordering of clitics according to their case, and also their differentiating the clitic forms themselves by case. Pashto falls outside this class of languages. Yet another unusual feature of Pashto is its dislocation of genitive clitics from the overt nominals with which they are semantically associated. A comparison of Pashto's possessor dislocation with similar constructions in other languages suggests that a novel treatment of this construction is warranted.

In previous research, the main difficulty in explaining Pashto 2P clitics has been to formulate rules that explain their placement without compromising independent principles of grammar that are assumed to hold robustly of other languages. Based on certain facts concerning the interaction of phonological processes and clitic placement, the principal claim of Tegey (1977) was that some phonological rules needed to apply before syntactic ones—a troubling conclusion for generative linguistics, which has long held that fully formed syntactic representations feed the phonological component. The need to reorganize the model of grammar in this way has been challenged from different perspectives by Kaisse (1981, 1985: 132–143) and Roberts (1997). The principal (and mistaken) claim of the latter was that 2P clitic placement was primarily a phonological effect, rather than a syntactic one, as both Tegey (1977) and Kaisse (1981, 1985: 132–143) had assumed.

This chapter will concur with earlier works that clitic placement is principally syntactic, but unlike earlier works, will be considerably more explicit in its formulation of the syntactic processes involved. In particular, the extent to which syntax determines clitic placement will be shown to be caused by the independent scrambling of nominals; the clitics themselves are never moved in the syntax. The first question to be addressed, then, which has not hitherto been asked, is whether 2P pronominal clitics are arguments (displaced from their base position in VP by a movement process) or agreement morphemes (of a kind similar to the agreement suffixes on the verb). Contra the implication of Tegey (1977), pronominal clitics will be shown to be agreement morphemes (rather than arguments) that are merged in positions that are high in the clause. This

conclusion results in a considerably simplified explanation of 2P clitic placement, with several striking results:

(i) **Clitics do not bear case.** Because clitics are agreement morphemes, only the actual argument (*pro*) that the clitic identifies bears case. Furthermore, there are not three sets of 2P pronominal clitics (ergative, accusative, and genitive), as it has been traditional to describe them. Rather, there is a single set of 2P pronominal clitics, which spell out person and number features under specifier-head agreement with *pro*, but *pro* itself bears a single (oblique) case.[32]

(ii) **The placement of clitics is independently resolved.** Because clitics are generated in fairly high positions (but below TP), there is no 'clitic placement rule' in the syntax; the only way in which clitics may be dislocated is by Prosodic Inversion (Halpern 1995), a 'Last Resort' PF option to save a structure in which a 2P clitic would lack a host. As far as PF is concerned, a clitic (or clitic cluster) can always be licensed in a structure; because every sentence contains a verb, a clitic may always take the verb as its host if a sentence does not contain any other overt material. The same is not true of licensing at LF, however: it is there that the argument structure of the verb ultimately determines whether the appearance of a pronominal clitic is licit.

(iii) **Ergative, accusative, and genitive arguments are checked by the same functional category.** This explains both the identity and complementarity of ergative and accusative clitics. A second agreement projection appears in clauses that contain a genitive clitic in addition to an ergative/accusative one, but the nature of this category is exactly the same:

---

[32] For clarity of exposition, it is often be useful to refer to individual clitics as ergative, accusative, or genitive, but these labels refer only to an individual clitic's actual or possible interpretation in a given sentence, which is derived from the base position of the *pro* that the clitic identifies.

it simply spells out person and number features, without reference to case, explaining why the genitive clitic has the same form as the ergative and accusative clitic (because its case is identical: oblique), and why ambiguity results in certain sentences having more than one clitic. In principle, as few or as many such agreement projections may be generated, but the argument structure of the verb (and the number of overt NPs that may take possessors) will ensure that all and only the required number of agreement projections are present.

(iv)   **Clitics and verbal agreement suffixes are identical in function.** 2P clitics identify oblique-case NPs (ergative, accusative, genitive), while verbal agreement suffixes identify the complementary direct-case NPs (nominative, absolutive).   The only difference between clitics and verbal agreement suffixes is their morphology: the former are merged as the heads of Agr projections above VP, while the latter are merged with the verb as a fully inflected word (hence may not be separated from their verbal stem). Nevertheless, by LF, the features associated with verbal suffixes are erased within the same system of Agr projections occupied by 2P clitics.

(v)   **Ordering within the clitic cluster is determined by syntactic structure.** There is no need for an arbitrary template to position the clitics with respect to each other, since clitics appear where the syntax has inserted them, as will be seen from the interpretive effects to be discussed in chapter 4. While there remain some clitic orders that may be best handled (at present) by the morphological or phonological component, even the problematic cases are nevertheless suggestive of more principled solutions.

The chapter is organized as follows. Section 3.2 provides an overview of 2P clitics and their placement. Section 3.3 reviews two recent analyses (van der Leeuw 1995, 1997; Roberts

1997) that do not adequately characterize clitic placement. Section 3.4 shows that 2P pronominal clitics are agreement morphemes, not arguments, and shows how the placement and interpretation of these morphemes is explainable by a single agreement projection. Section 3.5 turns to Pashto's genitive clitics, which are the most useful tool in establishing how second-position clitics function in the language. Finally, section 3.6 shows that the apparent clitic doubling that occurs in certain relative clauses is illusory, and that the relevant contrasts are straightforwardly explained, again, by regarding clitics as agreement morphemes that identify *pro*. In relative clauses and left-dislocation, this *pro* is resumptive (i.e., bound by an operator).

The approach taken here is therefore opposite to that taken earlier in Roberts (1997): where that paper pushed a phonological analysis of clitic placement as far as possible, this chapter will pursue a syntactic analysis as far as possible, which will be seen to yield far more satisfying results.

3.2. Overview of clitics and their placement

The second-position clitics of Pashto include pronominals, modals, and adverbials, listed below (Tegey 1977: 81):

(82)   Second-position clitics

Pronominal (ergative, accusative, genitive)

mee   1SG

dee   2SG

yee   3SG, 3PL

am   1PL, 2PL

mo   1PL, 2PL

<u>Modal</u>

ba   future, 'will', 'might', 'must', 'should', 'may'

dee   'should', 'had better', 'let'

<u>Adverbial</u>

kho   'indeed', 'really', 'of course'

no   'then'

The following paradigms illustrate that these clitics occur, informally speaking, in second position of the clause. As optional, sentence-initial items are removed, the clitics take as a host whatever other element appears initially. Here and throughout, the 2P clitics are underlined.

(83)   a.   kushal   <u>mee</u>   zyaati   ne   wah-i

               Khoshal   *1SG*   *anymore*   *NEG*   *hit-PRES3SG*

               'Khoshal does not hit me anymore'

     b.   zyaati   <u>mee</u>   ne   wah-i

               anymore   *1SG*   *NEG*   *hit-PRES3SG*

               'He doesn't hit me anymore'

     c.   ne   <u>mee</u>   wah-i

               NEG   *1SG*   *hit-PRES3SG*

               'He doesn't hit me'

     d.   wah-i   <u>mee</u>

               hit-PRES3SG   *1SG*

               'He hits me' (Tegey 1977: 132)

69

(84)   a.   tor   <u>dee</u>   nen   khar   ne   raawal-i

                Tor   *should*   *today*   *donkey*   NEG   *bring-*PRES3SG

                'Tor should not bring the donkey today'

    b.   nen   <u>dee</u>   khar   ne   raawal-i

                today   *should*   *donkey*   NEG   *bring-*PRES3SG

                'He should not bring the donkey today'

    c.   khar   <u>dee</u>   ne   raawal-i

                donkey   *should*   NEG   *bring-*PRES3SG

                'He should not bring the donkey'

    d.   ne   <u>dee</u>   raawal-i

                NEG   *should*   *bring-*PRES3SG

                'He should not bring it'

    e.   raawal-i   <u>dee</u>

                *bring-*PRES3SG   *should*

                'He should bring it' (Tegey 1977: 82–83)[33]

---

[33] In isolation, the (d) and (e) sentences are not grammatical, as they do not contain an object, either in the form of a nominal, or in the form of a clitic or verbal agreement. Jan Mohammad (p.c.) observes that these sentences are well-formed in the context of a paradigm, just as in English, one may say, for example, 'I love, you love, he loves' in order to illustrate verbal agreement, but without repeating an (irrelevant) grammatical object. Another interpretation of these sentences (Farooq Babrakzai, p.c.), which would render them grammatical (and illustrate the same pattern as above), would be to assume that they contained the 3sg accusative clitic *yee* (underlyingly /ee/), which would merge with the preceding 2sg clitic *dee* by a regular phonological process (about which, see sec. 4.5.3).

Pashto is fairly rigidly verb-final, and so (83d) and (84e) are of particular interest, as they illustrate that the clitic's need to have a host to its left is strong enough that it compels the verb to appear non-finally in a sentence containing only one word (the verb) other than the clitic. Although it is not obvious from the sentences above, clitics are positioned not after the first word, but rather after the first constituent, as the following examples make clear:

(85)   a.   [NP   aagha   sheel   kalena   danga   aw   khaaysta   peeghla]

           that   20   year   tall   and   pretty   girl

           <u>mee</u>   nen   byaa   welida

           I   today   again   saw

           'I saw that twenty-year-old tall and pretty girl again today'

       b.   [NP   khushal   aw   patang]   <u>ba</u>   <u>yee</u>   der   ta   raawrri

           Khosal   and   Patang   will   it   you   to   bring

           'Khosal and Patang will bring it to you' (Tegey 1977: 83–84)

The clitics in (82) may co-occur—as illustrated by such sentences as (85b)—and when they do, their respective order is fixed, as schematized by the following template (Tegey 1977: 191).

(86)   *1*   *2*   *3*   *4*
           kho   ba   am   am
                                                  mo
           'indeed'   'will'   1PL, 2PL   1PL, 2PL

           5   *6*   *7*   *8*
           mee   dee   yee   no
           1SG   2SG;   3SG, 3PL   'then'
                        'should'

The modal 'should' and the 2sg pronoun are homophonous *(dee)*, have the same position with respect to surrounding clitics, and do not co-occur, hence they are listed in a single column.

71

Column 4 contains two forms, a discussion of which will be deferred until the following chapter, in which this template will be discarded, as its effects are derivable from syntax.

3.3. Previous analyses

The first analyses of Pashto 2P clitics were Tegey (1977) and Kaisse (1981, 1985: 132–143), both of whom assumed that clitic placement was a syntactic process, though without making its details explicit. Tegey (1977: 122), for example, suggested that 2P clitics "are placed after the first major surface constituent that bears at least one main stress—where 'major constituent' may be directly dominated by S, VP, or V." As this statement conflates syntactic and phonological processes, it cannot be regarded as a syntactic rule (i.e., as a process that occurs before Spell Out).

Tegey's disjunctive statement regarding S, VP, and V, was necessary because his analysis predated a syntactic framework in which empty pronominals were recognized as constituents in their own right. The disjunction was remedied by Mohammad (1993: sec. 2), who offered the following rule:

(87) Place the 2P clitics after the first phonologically realized syntactic constituent of the following form within S, [$_{XP}$ ... Y ...], where Y=N, V, Adj, Adv. If V is compound, place the clitic after the stress-bearing constituent of the V.

A difficulty with this rule is, again, that it conflates syntactic and phonological processes in a seemingly unconstrained fashion. A more satisfactory solution to 2P clitic placement is suggested by Mohammad (1993: sec. 3) and Babrakzai (1999: ch. 3), who attempt to correlate the appearance (or absence) of clitics with agreement. This analysis will be pursued in detail in the following section, where it will be shown to explain not only the occurrence of 2P clitics, but also their interpretation and placement.

Two phonological approaches to 2P clitic placement have also recently been offered (van der Leeuw 1995, 1997: sec. 5.1; Roberts 1997), but neither of them is satisfactory. The analysis of van der Leeuw focuses on separable prefix forms of the kind in (88) below. While the initial *a* of the verb *akhistel* 'to buy, to take' may have been a prefix historically, neither *a* nor *khistel* has independent synchronic meaning. In clauses containing only a verb in addition to the clitic—as in the conditional clause below—the clitic follows the initial, stressed *a* (which often becomes back *aa*, a variant noted also by Tegey 1977: 179 n. 2). The clitic may not follow the verb in such cases, but rather must divide the parts of the verb:

(88) áa   mi   khiste   kho   hágha   wrost   wo

      PREFIX   1SG   buy   but   that   rotten be

      'I would have bought it but it was rotten' (Kandahar)

Van der Leeuw therefore suggests the following constraint to explain clitic placement:

(89)   ALIGN (affix, L, syllable', R)

          Align the left edge of the (postlexical) affix [=clitic] to the right edge of the stressed

          syllable, where *syllable'* means stressed syllable. (van der Leeuw 1997: 130)

This analysis suffers from several problems. First, the clitic does not simply follow a stressed initial syllable, but more specifically follows a (historic) prefix. Second, the constraint treats 2P clitics (Tegey's "Group I" clitics) as though they were verbal clitics. Pashto has a separate set of verbal clitics—Tegey's (1977: ch. 7) "Group II" clitics, which indicate dative roles, and which do not intervene between the morphemes of a verb (Tegey 1977: 240). It is crucial, therefore, to distinguish the two groups. While the constraint in (89) accounts for clitic placement in sentences like (88), merely adding an adverb to the beginning of such a sentence will coax the clitic away from the verb, and into second position:

(90)　　　parun　　mi　　áakhiste　kho　　hágha　　wrost　　wo

　　　　　　yesterday　*1SG*　*buy*　　　*but*　　*that*　　*rotten*　　*be*

　　　　'Yesterday I would have bought it but it was rotten' (Kandahar)

(91)　　　*parun　　áa　　mi　　khiste　　kho　　hágha　　wrost　　wo

　　　　　　yesterday　*PREFIX*　*1SG*　*buy*　　*but*　　*that*　　*rotten*　　*be*

As adjuncts are added to the beginning of the sentence, the clitic may appear quite far away from the verb:

(92)　　　parun　　mi　pe　maaket　ki　áakhiste　kho　　hágha　wrost　wo

　　　　　　yesterday　*1SG*　*at*　*market*　*at*　*buy*　　　*but*　　*that*　*rotten*　*be*

　　　　'Yesterday at the market I would have bought it but it was rotten' (Kandahar)

(93)　　　*parun　　pe　maaket　ki　mi　áakhiste　kho　　hágha　wrost　wo

　　　　　　yesterday　*at*　*market*　*at*　*1SG*　*buy*　　*but*　　*that*　*rotten*　*be*

The same point may be demonstrated for the perfective morpheme *waa* (its different form here—
waa vs. *we*—will be discussed in chapter 4; see (303) below), which has often been regarded as a
prefix, but which is regarded here as a proclitic morpheme heading AspP, as per the discussion in
chapter 2:

(94)　　　a.　　wáa　　dee　　khist-el

　　　　　　　　　PERF　　*2SG*　　*buy-PAST(MASC PL)*

　　　　　　　'You bought them'

74

b.   parun   <u>dee</u>   wáa   khist-el

      yesterday   2SG   PERF   *buy-PAST(MASC PL)*

      'You bought them yesterday' (Laghman)

c.   *parun   wáa   <u>dee</u>   khist-el

      yesterday   PERF   2SG   *buy-PAST(MASC PL)*

      'You bought them yesterday'

Clearly, the constraint in (89) will not place clitics in the correct position in sentences that contain anything more than a single verb. Aside from its empirical inadequacy, the constraint in (89) faces the familiar theoretical difficulty: it conflates syntactic and prosodic requirements by referring to stress and morpheme structure in a single statement.

The analysis of Roberts (1997) sought to separate the syntactic and phonological properties of 2P clitic placement by referring instead to properties of prosodic structure. In this approach, syntactic categories (specifically, maximal projections) induced Phonological Phrase boundaries (Selkirk 1984, 1986). By regarding clitic placement as a constraint that aligns clitics to those boundaries, as in (95), the syntactic and phonological properties of 2P were separated:

(95)   ALIGN (cl, L, PPh, R)

       Align the left edge of a clitic to the right edge of a phonological phrase.

Unfortunately, this account had little to say about why a clitic could separate a prefix from a verb stem, beyond stipulating that such prefixes as occur in (88) were exceptional in inducing their own Phonological Phrase boundary. Inasmuch as it was unprecedented for a single vowel to constitute its own Phonological Phrase, this account remained unsatisfactory.

Another problem with this analysis betrayed the syntactic properties of 2P clitic placement: the clitics do not intervene among conjuncts, nor indeed among the parts of any clause-initial constituent, as was also illustrated above in (85):

(96)   a.   [_{ConjP} pekol   aw   koteey]   <u>mi</u>   ághoste   kho ...

                     cap   *and*   *jacket*   *1SG*   *wear*   *but*

             'I would've worn the cap and jacket but ...' (Kandahar)

      b.   *pekol   <u>mi</u>   aw   koteey   ághoste ...

           cap   *1SG*   *and*   *jacket*   *wear*

      c.   *pekol   aw   <u>mi</u>   koteey   ághoste ...

           cap   *and*   *1SG*   *jacket*   *wear*

In order to maintain the constraint in (95), it was necessary to suggest that coordinate structures were not represented hierarchically or linearly, but rather were parallel structures of the kind suggested by Williams (1977, 1978) and Goodall (1987), such that coordinate structures induced a single Phonological Phrase boundary. Such a move was unprecedented and unusual, in that a non-linear representation was determining the structure of a level of representation (PF) that is crucially linear.

Despite ample such difficulties, the syntactic properties of 2P clitics remain to be examined. The research plan has thus returned to the original assumption of Tegey (1977) and Kaisse (1981, 1985: 132–143) that clitic placement is syntactic, although now with greater attention to ensuring that syntactic and phonological processes are treated separately. The first two chapters having established a basic phrase structure for the language, we are now in a position to explore the role of clitics within this structure. It will be seen that a syntactic

approach to 2P clitics can explain their placement and interpretation without the methodological and theoretical difficulties faced by earlier accounts.

3.4. Clitics as agreement

As reviewed in the previous section, attempts to explain the placement of 2P clitics in Pashto have broadly assumed either that the rule of clitic placement is syntactic (Tegey 1977; Kaisse 1981, 1985: 132–143) or phonological (van der Leeuw 1995, 1997; Roberts 1997). However, none of these analyses, including the syntactic ones, has considered the base position of clitics. Although Tegey usually seems to assume that clitics move to their surface position, he is never explicit about where they move from.[34]

The issues that remain to be addressed are (i) whether pronominal clitics represent arguments or agreement morphemes; (ii) their base position; (iii) their landing site; (iv) their internal ordering; and (v) if clitics do indeed move from their base position, whether they move by a phonological or syntactic process, or a combination of both. This section addresses the first question, adducing evidence that clitics represent properties of agreement (Suñer 1988, Franco 1991, 1993, Davis 1995, 1998, Sportiche 1996, Rudin 1997), rather than being the actual arguments of the verb (Jelinek 1984, Kayne 1989b, Baker 1996, Progovac 1999). It will be shown that once second-position clitics are treated as agreement morphemes, their interpretation and positioning within the clause may be derived from independent principles of syntax and phonology. The evidence to be considered includes word order, the distribution of clitics with respect to verbal agreement, and coordination.

---

[34] A typical comment is the following: "In order to distinguish the Type-I clitics, which move to second position, from the Type-II clitics, which move to the verb, the clitic movement rules would *also* have to refer to syntactic function" (Tegey 1977: 242). Such statements imply that Tegey regarded the clitics as arguments of the verb, although he cannot be held too strongly to having made such a claim, as he never explicitly made it.

3.4.1. Deriving the second position effect

The interpretation, positioning, and internal ordering of clitics are explained by more than one component of grammar, with syntax playing a larger role than has hitherto been suggested. The interpretation of clitics is best explained by the verb's argument structure (stated as conditions holding at LF), while their appearance in second position is best explained by their appearing in functional (agreement) projections that are merged in high positions.[35] Because of the availability of subject and object *pro*-drop, it is possible for the syntax to derive a structure in which the clitics do not have a suitable phonological host to their left; in such cases, a Last Resort option of Prosodic Inversion applies at PF. There are other phonological constraints that apply at PF to repair illicit sequences of clitics; such repair strategies always occur as a last resort, though, and do not have any interpretative import. Derivations must converge at both LF and PF, and the remainder of this chapter and the next will suggest how the various facets of licensing a clitic is divided between LF and PF. The necessary constraints can be stated with considerable generality, allowing many ill-formed derivations to be straightforwardly excluded.

Clitics in various languages have been argued to have properties of agreement and/or arguments, and for this reason they have been analyzed as involving either base-generation or movement. Sportiche (1996) seeks to unite the two approaches by suggesting that each pronominal clitic heads its own projection, called here Clitic Phrase.[36] In languages without clitic doubling, arguments are assumed to be the empty category *pro*, which moves to Spec/CliticP to check its features (e.g., case, person, number). In languages with clitic doubling,

---

[35] Several analyses of second-position clitics in Serbo-Croatian have assumed that the clitics cluster in C_0 or some other high projection (Progovac 1993, 1999; Schütze 1994; Tomic 1996). The present analysis of Pashto therefore bears more resemblance to Stjepanovic's (1998) treatment of Serbo-Croatian—and Davis' (1996, 1998) treatment of St'át'imcets—in which each second-position clitic occupies a distinct functional projection external to VP. Overviews of recent literature on such issues are offered by Rudin (1997) and Franks (2000).

[36] Sportiche (1996: 235) calls these projections 'clitic voices', and assigns them labels like 'Nom Voice', 'Accusative Voice', and 'Dative Voice'. Such distinctions are not made in the category labels here, since pronominal clitics in Pashto bear a single (oblique) case, and their interpretation is ambiguous, as will be seen.

the overt argument moves to Spec/CliticP, also to check its features. The canonical configuration

for clitics is therefore as follows, with Spec/CliticP being either the empty argument *pro,* or an

overt argument (XP), as in a clitic-doubling language.

(97)   CliticP

            pro/XP   Clitic'
                          |
                        Clitic
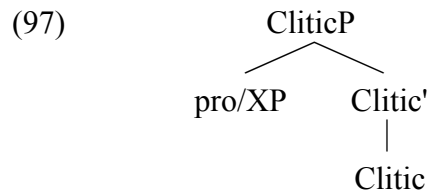
Languages vary as to whether movement of the argument to Spec/CliticP occurs before or after

Spell Out, although it must happen by LF:

(98)   Clitic Criterion (Sportiche 1996: 236)

              i.   A clitic must be in a Spec-head relationship with a [+F] XP at LF

              ii.   A [+F] XP must be in a Spec-head relationship with a clitic at LF

The parameters of variation pertain to whether categories and/or movement are overt or covert

(Sportiche 1996: 237):

(99)   Clitic construction parameters

              i.   The clitic head is overt or covert

              ii.   The argument XP is overt or covert

              iii. Movement of the argument XP to Spec/CliticP occurs overtly or covertly

This proposal was intended by Sportiche to explain Romance pronominal clitics and accusative

scrambling in Germanic. Although he explains at length that any analysis of pronominal clitics

must apply cross-linguistically, he does not consider whether 2P clitics may be treated in the same fashion. It will be shown now that this analysis may indeed be extended to 2P clitics.

The first step in extending this analysis to 2P clitics is to note that the label 'CliticP' in (97) is inaccurate, inasmuch as it seems to identify clitics as a natural class. As Halpern (1995) and others have noted, 'clitic' is not a uniform notion, covering as it does a variety of phenomena. The notion of clitic will similarly have no independent theoretical status here. As the structure in (97) is appropriate only for pronominal clitics—more specifically, features of person, number, and case—it is best regarded as a projection of those features. The category label is irrelevant, and in precursors to the present framework, equivalently could have been labeled AgrS, AgrO, AgrIO, etc. The label CliticP will sometimes be retained in the following discussion for clarity, though it should be remembered that CliticP is simply an agreement projection. Auxiliary and modal 2P clitics occur in a different configuration altogether (as might be expected from their different features and interpretations), and these will be explained in the following chapter. This chapter focuses only on the 2P pronominal clitics.

The second step in extending this analysis to 2P pronominal clitics is to dispense with the Clitic Criterion in (98). The Clitic Criterion may be stated more generally, perhaps as a Generalized Licensing Criterion (Sportiche 1996: 264), but surely more accurately as nothing more than the regular specifier-head agreement that must obtain between features in order for a derivation to converge. In this respect, the clitic construction parameters in (99) are also descriptive artifacts, the actual parametric effects being derived from the properties of features that are drawn from the lexicon: their having an overt PF form or not, and their being strong or weak, will suffice to derive the parameters in (99). That Pashto realizes both options of (99ii)— in having clitic doubling with dative verbal clitics, and no clitic doubling with the ergative, accusative, and genitive 2P clitics—is ample evidence that the parameter does not have any independent status.

Turning to concrete examples, the principal questions to be addressed for 2P pronominal clitics are (i) whether CliticP is head-initial or head-final; and (ii) where it is located in phrase

structure. Regarding (i), the clitic head is a functional category, carrying canonical features of agreement, viz., person and number, which are checked by the assorted pronominal clitics; and like other functional categories, CliticP is head-initial. Because CliticP is a functional projection, well above the *v*P-internal arguments, and because the verb is rigidly in final position, CliticP could not be head-final, as it would result (incorrectly) in clitics following the verb. Regarding question (ii), as an agreement projection, CliticP is located in the expected position, lower than the complementizer (within TP), but higher than the perfective and negative morphemes. This may be seen from the position of the 1sg clitic *mee* in the embedded clause of (100):

(100)   tor   mee   we   lid-e

         Tor   *1SG*   *PERF*   *see-MASC3SG*

         [CP magar [TP spin   mee   we   ne   lid-e]]

            but   *Spin*   *1SG*   *PERF*   *NEG*   *see-MASC3SG*

         'I saw Tor, but I didn't see Spin' (Tegey 1977: 127)

Recall from chapters 1 and 2 that the clause has the following structure before movement of arguments:

(101)   [TP ... [AspP ... [NegP ... [*v*P EXT ARGUMENT [VP INT ARGUMENTS ...]]]]]]

Because CliticP is lower than TP, but higher than AspP, most second position effects are straightforwardly derived by the syntax. The embedded sentence of (100) has the following structure at Spell Out:

(102)

```
                    TP
            ┌───────┴───────┐
           NP            CliticP
            |          ┌────┴─────┐
         Spinᵢ       Clitic'
                   ┌─────┴──────┐
                 Clitic        AspP
                   |       ┌─────┴─────┐
               mee 1SG   NegP         Asp
                       ┌───┴────┐      |
                     Neg       vP   we- PERF
                      |     ┌───┴───┐
                     ne    NP      VP
                           |      ╱──╲
                          pro   tᵢ lide 'see (MASC3SG)'
```
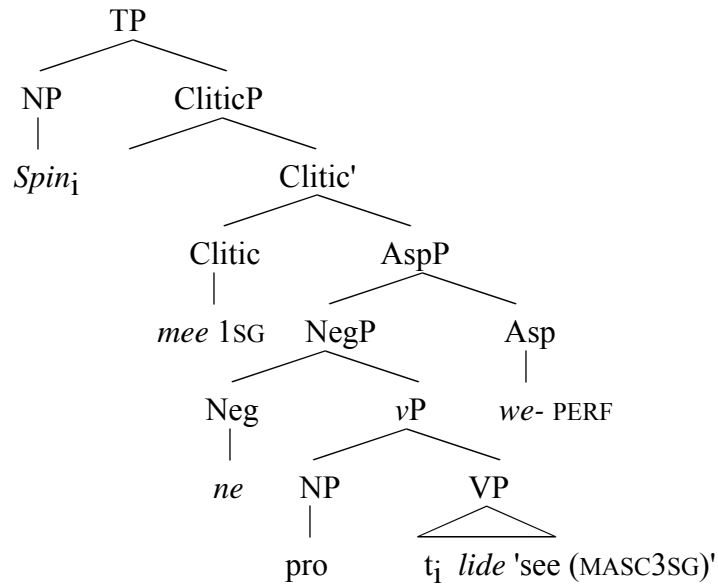
The object has moved to initial position, because in the absence of an overt, full NP subject, the object may move to Spec/TP to check the EPP feature. (See the following section.) The trace is interpreted as an object, due to its structural position with respect to the verb, with which it also agrees (due to past-tense ergativity). Because this is a transitive verb, it projects an external argument, *pro*, in Spec/*v*P. As this argument must be interpreted, and because the person and number features in CliticP must be erased, *pro* moves to Spec/CliticP by LF:

(103)

```
                    TP
            ┌───────┴───────┐
           NP            CliticP
            |        ┌──────┼────────┐
         Spinᵢ     proₖ         Clitic'
                             ┌─────┴──────┐
                           Clitic        AspP
                             |       ┌─────┴─────┐
                         mee 1SG   NegP         Asp
                                 ┌───┴────┐      |
                               Neg       vP   we- PERF
                                |     ┌───┴───┐
                               ne    NP      VP
                                     |      ╱──╲
                                    tₖ    tᵢ lide 'see (MASC3SG)'
```
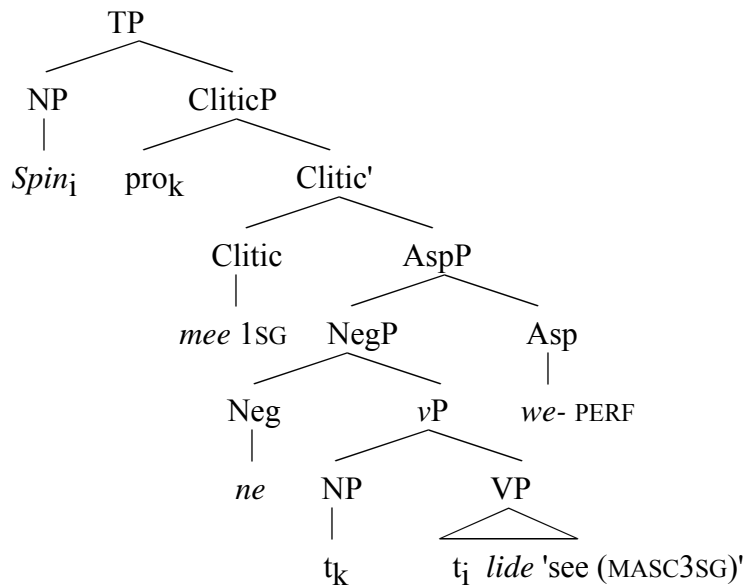
The relation thus established serves to identify the *pro* argument as 1sg. If CliticP were absent, the derivation would crash at LF, as the external argument *pro* could not be identified. That the clitic is interpreted as the subject is due to its agreeing with *pro*, which has originated as the external argument. There is no need for the clitic to bear a distinct 'ergative case' feature, its interpretation as the subject being clear from the base position of *pro*—a desirable result, since clitics having this form may also (and often ambiguously) function as objects and possessors, as will be shown. Evidence that movement of *pro* to check person and number features is deferred until LF is that movement of the object (which can be seen to occur in overt syntax) before the subject would be a violation of cyclicity if the subject (*pro*) moved in overt syntax. Because cyclicity does not constrain LF operations (Chomsky 1995: 255; Kitahara 1995: 56), and because movement of *pro* (as an empty category) need happen only in order to be interpreted at LF, by Procrastinate it may be deferred until after Spell Out.

While considerations like the above are independently needed in any analysis of clitics in which they are agreement morphemes, it has the pleasing result that it escapes a potential 'look-ahead' problem by deriving the surface position of the clitic without movement of the clitic in order to satisfy phonological conditions regarding its host. In (103), the surface position of the 2P pronominal clitic *mee* has been derived entirely by the syntax; it is not in any sort of 'second position' with any independent status, but is simply heading an agreement projection between TP and AspP. While the clitic does indeed require a phonological host to its left, this requirement is vacuously satisfied due to the independent movement of the object. Evidence for this analysis of clitics as agreement of person and number features will now be presented.

3.4.2. Scrambling
3.4.2.1. EPP and focus as triggers

Scrambling in embedded clauses offers compelling evidence for the analysis sketched above, offering insight into (i) the surface position of overt NPs, and (ii) the base position of second-

position clitics. The EPP (a requirement that Spec/TP be filled at Spell Out) is argued to be the driving force behind the placement of second-position clitics, as long as pronominal clitics are regarded as heading the clause's agreement projection.

It is assumed that scrambling is not optional movement; rather, the principal forces that drive scrambling are features involving case, agreement, focus, and so on (Chomsky 1995; Miyagawa 1997; Boškovic and Takahashi 1998). Miyagawa suggests more specifically that A-scrambling is driven by some feature on T(ense)—either case or agreement (Miyagawa 1997) or the EPP (Miyagawa to appear)—while A'-scrambling is driven by a focus feature. Japanese, like Pashto, is a (lexical) head-final language with scrambling; both SOV and OSV word orders are available, but the former is taken to be the default order, while the latter is derived by fronting the object:

(104) a. Taroo-ga piza-o   tabeta

      Taro-NOM *pizza-ACC* *ate*

      'Taro ate pizza'

    b. Piza-o   Taroo-ga tabeta

      pizza-ACC *Taro-NOM* *ate*

      'Taro ate pizza' (Miyagawa to appear: §2)

By considering the scope of the quantifier *zen'in* 'all' (in subject versus object position) with respect to negation, it can be shown that the apparently free order exemplified by (104) is in fact constrained. When 'all' is in object position, it must be interpreted inside the scope of negation; when 'all' is in subject position, the scope is reversed:[37]

---

[37] The following Japanese sentences are presented in Miyagawa (to appear) as subordinate clauses, selected by either *yo* (excl) or *to omou* '(I) think that...', but these items are omitted here for brevity.

(105)   a.   Taroo-ga   zen'in-o   home-nakat-ta

                  Taro-NOM   all-ACC   praise-NEG-PAST

                  'Taro didn't praise all' (not >> all; *all >> not)

(b.)   Zen'in-ga   sono   tesuto-o   uke-nakat-ta

                  all-NOM   that   test-ACC   take-NEG-PAST

                  'All did not take that test' (all >> not; *not >> all)

                                       (Miyagawa to appear: §2)

This subject-object asymmetry may be explained if negation occurs above *v*P, c-commanding the quantificational object in its VP-internal base position. The position of negation in Japanese is therefore the same as in Pashto—recall (69) in chapter 2—except that NegP and TP are head-final in Japanese:

(106)
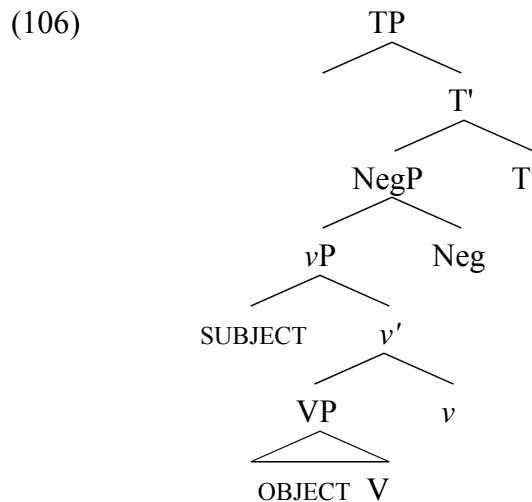
This structure explains (105a), since the quantifier is in object position, c-commanded by negation. In order to explain (105b), in which the quantifier is in subject position (merged at Spec/*v*P), it may be assumed that the subject has moved to Spec/TP. The exact trigger for this

movement, though, is not yet clear: it could be the EPP, or a strong nominative feature. In any case, once the subject quantifier is in Spec/TP, it c-commands negation, as desired, taking wide scope overt it. Evidence that it is specifically the EPP that drives object-scrambling (rather than a strong nominative feature) may be seen from the shifted scope that accompanies the OSV variant of (105b):

(107)   sono   tesuto-o   Zen'in-ga   uke-nakat-ta

          that   test-ACC   all-NOM   take-NEG-PAST

          'All did not take that test' (not >> all; all >> not)   (Miyagawa to appear: §2)

The wide-scope reading for negation that was unavailable in the SOV order in (105b) becomes available in the OSV variant in (107), suggesting that the quantificational subject remains *in situ*, in Spec/*v*P, under the scope of negation. Since the subject can be seen here not to move obligatorily to Spec/TP, nominative must *not* be a strong feature; rather, movement to Spec/TP is driven only by the strong EPP feature there, which may be erased either by the subject, as in (105b), or by the object, as in (107). In order for the object to move to Spec/TP in lieu of the subject, however, the subject and object must be equidistant from T; Miyagawa suggests that verb-to-T head-movement gives this result. The structure of (107) would therefore be as follows:

(108)

```
                    TP
           ┌────────┴────────┐
       OBJECT_i            T'
               ┌───────────┴───────────┐
             NegP                     T
        ┌─────┴─────┐           ┌──────┴──────┐
       vP          t_X        V-v-Neg_X-T
   ┌────┴────┐
SUBJECT      v'
        ┌─────┴─────┐
       VP          t_X
    ┌───┴───┐
   t_i     t_X
```
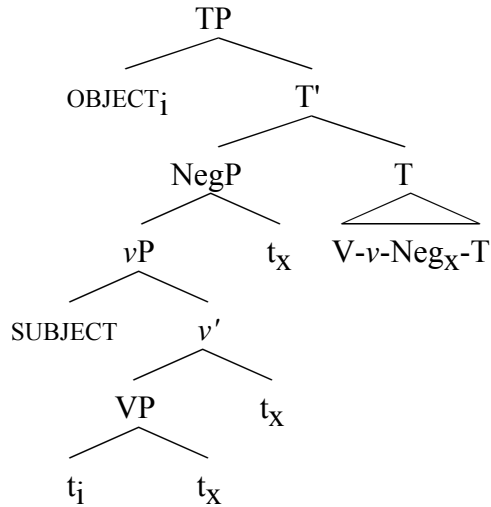
<div align="right">(Miyagawa to appear: §2)</div>

As (107) is ambiguous, it remains to explain the second interpretation, in which 'all' has wide scope. Because both the subject and object are equidistant to T, either argument may move there in order to satisfy the EPP. The 'all >> not' reading may be derived by assuming that the subject has moved to Spec/TP, satisfying the EPP, and that the object undergoes A'-movement for focus. The structure would be as follows:[38]

(109)

```
                 FocusP
           ┌───────┴───────┐
       OBJECT_i           TP
                  ┌────────┴────────┐
              SUBJECT_k            T'
                      ┌────────────┴────────────┐
                    NegP                       T
               ┌─────┴─────┐            ┌───────┴───────┐
              vP          t_X         V-v-Neg_X-T
          ┌────┴────┐
         t_k        v'
               ┌─────┴─────┐
              VP          t_X
           ┌───┴───┐
          t_i     t_X
```

---
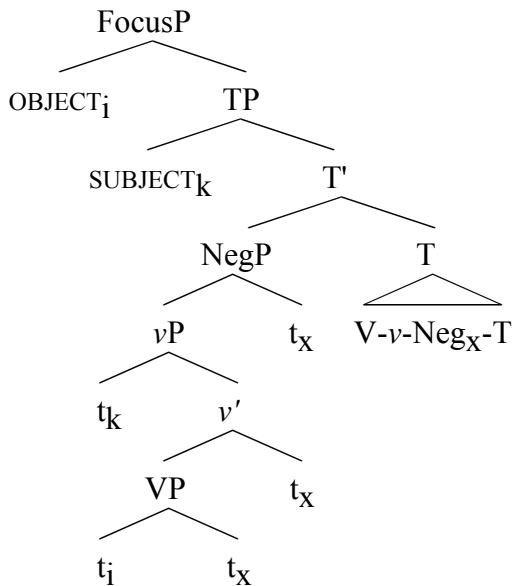
[38] See Cinque (1999: 225 n. 25) and the references there for suggestions that the CP projection should be articulated to include projections for Topic, Focus, etc.

In this structure, the quantificational subject c-commands negation, deriving the wide-scope reading for 'all', exactly as desired. Further evidence that the two interpretations of (107) have the different representations in (108–109) is that such sentences may be disambiguated with an appropriate adverb. The temporal adverb *kinoo* 'yesterday', for example, occurs in the projection of T, and so when the subject appears to the left of the adverb—as it does in (110)—the subject must indeed be in Spec/TP, while the object is in the focus position.

(110)　　　Kono　ronbun-o　　zen'in-ga　kinoo　　　yoma-nakat-ta

　　　　　　this　　article-ACC　all-NOM　　yesterday　read-NEG-PAST

　　　　　　'This article, all did not read yesterday' (all >> not; *not >> all)

　　　　　　　　　　　　　　　　　　　　　　　　(Miyagawa to appear: §2)

This analysis implies that scrambling is highly constrained as follows (Miyagawa to appear):

(111)　　　Triggers for scrambling

　　　　　　A-scrambling:　EPP feature on T

　　　　　　A'-scrambling:　Focus

The next section examines how Pashto behaves with respect to these triggers.

3.4.2.2. Last resort

The EPP is likewise a strong feature in Pashto, and the fact that it compels either the subject or object to move overtly will be shown to have the effect of placing clitics in 'second position'. Turning first to word order effects, as shown in (112), the 3sg clitic *yee* may follow the complementizer *tshee*, and may also minimally invert with the verb. The clitic may not, however, move farther to the right, as shown in (112c):

(112) a. hágha dzhorra tshee <u>yee</u> aghust-ay wa genda wa

DET *suit* *COMP* *3SG* *wear-part* *was* *dirty* *was*

'the suit that he was wearing was dirty'

b. hágha dzhorra tshee aghust-ay <u>yee</u> wa genda wa

DET *suit* *COMP* *wear-part* *3SG* *was* *dirty* *was*

'the suit that he was wearing was dirty'

c. *hágha dzhorra tshee aghust-ay wa <u>yee</u> genda wa

DET *suit* *COMP* *wear-part* *was* *3SG* *dirty* *was*

'the suit that he was wearing was dirty'

(Jan Mohammad, p.c.)

The position of the 3sg clitic *yee* in (112a), immediately following the complementizer, presumably represents its base position, heading AgrP. Since the embedded clause lacks an overt NP that could have scrambled to the embedded Spec/TP (where the NP would then precede the clitic), the clitic may remain *in situ*, taking the complementizer *tshee* as its phonological host. The following pair of sentences demonstrates the same point:

(113) a. kéla tshi <u>mee</u> welid-e, naadzhórr(a) wu

when *COMP* *1SG* *saw-MASC3SG* *sick* *be(PAST.IMPF3SG)*

'when I saw him, he was sick'

b. *kéla tshi welid-e <u>mee</u>, naadzhórr(a) wu

when *COMP* *saw-MASC3SG* *1SG* *sick* *be(PAST.IMPF3SG)*

'when I saw him, he was sick' (Jan Mohammad, p.c.)

As agreement morphology on the past tense verb *welide* 'saw' in (113a) identifies the null object, the 1sg clitic *mee* saturates the verb's argument structure by identifying the (ergative) *pro* subject. There being no full NPs in the clause, the 1sg clitic *mee* may remain in its base position, heading AgrP), taking the complementizer *tshi* as its phonological host. There is no need for the clitic to undergo prosodic inversion, and indeed ungrammaticality results if it does so, as shown in (113b). It is not clear why prosodic inversion is ungrammatical in (113b), but possible in (112b).[39] This variation is irrelevant, however, as the only important point about these sentences is illustrated by (112a) and (113a): that the clitic may immediately follow the complementizer, in contrast to the sentences to be given below.

The placement of the clitic in sentences like (112) and (113) contrasts with its placement in sentences like the following, in which the subordinate clause containing the clitic also contains a full NP. In such sentences, the clitic may not follow the complementizer.

(114)   a.   agha   kisa   tshee   dzhan   yee[40] lwel-i   ddeera ugda da

                  DET   *story* *COMP*   *John*   *3SG*   *read-PRES3SG*   *very*   *long* *be*

                  'the story that John is reading is very long'

           b.   *agha   kisa   tshee   yee   dzhan   lwel-i   ddeera ugda da

                  DET   *story* *COMP*   *3SG*   *John*   *read-PRES3SG*   *very*   *long* *be*

                  'the story that John is reading is very long' (Jan Mohammad, p.c.)

---

[39] According to Babrakzai (1999: 49, 70 fn. 4), an ergative clitic may follow the complementizer in literary Pashto. Perhaps the availability of more than one location for clitics in some sentences betrays similar effects of style or register.

[40] Omitting the 3sg clitic *yee* from this relative clause does not result in ungrammaticality, but it becomes marginal; it is preferred to have the clitic at the site of relativization (Jan Mohammad, p.c.). The role of clitics inside relative clauses is discussed in sec. 3.6.

The contrast between (112a) and (113a) versus (114b) shows that a second-position clitic may immediately follow the complementizer *tshee/tshi* only when the clause does not contain a full NP. This contrast is taken as evidence that the embedded subject, 'John', moves to Spec/TP (to erase the EPP feature), where it precedes the agreement projection headed by the 3sg clitic *yee*.

Strikingly, this paradigm also obtains in sentences having only an overt object NP, suggesting that an overt object NP also moves to Spec/TP to check the EPP feature. In the embedded clause of (115a) below, the direct object—the 2sg strong pronoun *te*—immediately follows the complementizer *tshi*, just as the overt subject 'John' does in (114a). The 3sg clitic *yee* then follows the direct object:

(115)   a.   hágha   sarray   [tshi   te   <u>yee</u>   milma   kérr-ey   wee]   laarr

                that   *man(DIR)*   *COMP*   *PN2SG(DIR)*   *3SG*   *guest*   *do-PART*   *was*   *go*

                'the man who had invited you left'

    b.   ?hágha   sarray   [tshi   <u>yee</u>   te   milma   kérr-ey   wee]   laarr

                that   *man(DIR)*   *COMP*   *3SG*   *PN2SG(DIR)*   *guest*   *do-PART*   *was*   *go*

                'the man who had invited you left'

    c.   *hágha   sarray   [tshi   te   milma   <u>yee</u>   kérr-ey   wee]   laarr

                that   *man(DIR)*   *COMP*   *PN2SG(DIR)*   *guest*   *3SG*   *do-PART*   *was*   *go*

                'the man who had invited you left' (Jan Mohammad, p.c.)

These sentences suggest that either an overt subject or an overt object may move to Spec/TP in order to satisfy the EPP, as Miyagawa (to appear) has argued for Japanese.[41] If the direct object

_____

[41] Recall from ch. 1 that psych-predicates require non-direct (non-nominative) NPs as subjects, and so there are no case-related reasons that might be expected to exclude non-subject NPs from moving to Spec/TP.

te does not move, as in (115b), the sentence becomes marginal. Notice, specifically, that the less-than-perfect status of (115b) cannot be due to the 3sg clitic *yee* taking the complementizer *tshi* as its host, as this configuration was seen to be grammatical in (112a) and (113a). The difference between the embedded clauses of (112a) and (113a) versus those in (114a) and (115a) is that the former do not contain any full NPs, whereas the latter do. In the derivation of all of these sentences, the clitic remains *in situ*, heading AgrP. In the embedded clauses of (112a) and (113a), the clitic may therefore take the immediately preceding complementizer as its host. In the embedded clauses of (114a) and (115a), there is an overt NP in addition to the clitic, and the fact that the clitic follows the overt NP in these sentences shows not only that the overt NP moves before Spell Out, but that both subjects (114a) and objects (115a) undergo this movement. Furthermore, these data support a recent idea that constituents satisfying the EPP feature of T must also be in an agreement relation with some feature on T. The subject 'John' in (114a) is in Spec/TP and agrees with verb *lwel-i* 'read(PRES3SG)'; likewise, the 2sg pronominal object in (115a)—also in Spec/TP to satisfy the EPP—agrees with the verb *wee* 'be(2SG PAST IMPF)'. In sentences lacking overt arguments, the EC *pro* satisfies the EPP. When such sentences are past tense, then, as in (112–113) above, it is presumably the object *pro* (with which the verb agrees) that moves overtly to Spec/TP to satisfy the EPP, while the subject *pro* moves to Spec/CliticP in order to be identified by the ergative clitic.

Entirely in line with material presented earlier, then, these data suggest that second-position clitics do not move to a special 'second position' of the clause, or even to C^0, since no single definition of second position is possible. By merging the clitics as head of AgrP, however, their position is independently determined by whatever overt NPs move in the sentence. Since EPP is a strong feature, an NP (either the subject or the object) must move to satisfy it. When the NP is covert (*pro*), its movement to Spec/TP will not induce word order effects with respect to the clitic. But when the NP is overt, the second-position clitic will inevitably appear to the right of the NP, since TP selects AgrP. The placement of second-position clitics is thus reduced to independently required principles.

The most obvious alternative analysis—to assume that NPs do not move, but that pronominal clitics are arguments that move into a special clitic position—is untenable. If the clitic were moving to satisfy prosodic constraints (such as to find a suitable host), we would not expect a contrast between sentences like (112a) and (113a)—in which the clitic is prosodically hosted by the complementizer—versus those in (114) and (115), in which the clitic may not follow the complementizer. Note moreover that it is impossible to appeal to any sort of subject/object asymmetry, since none exists in these paradigms; the embedded clause of (114a) shows SOV order, while the embedded clause of (115a) shows OSV order, and so it cannot be the case that the clitic is moving, but rather that the single, overt NP is moving in both of these sentences. Perhaps a prosodic analysis could be salvaged by supposing that the clitic prefers to take a full NP as its host, if one is present, but otherwise need not, by an Optimality-Theoretic ranking of constraints accompanied by an appropriate definition of heaviness. But given the difficulty in formulating such constraints, the syntactic account proposed here—which is quite simple—will be maintained: both subject and object NPs are able to satisfy the strong EPP feature of the clause.

If both the subject and object of a clause are represented by full NPs, it is the subject that checks the EPP feature in Spec/TP, as might be expected by Shortest Move (since the subject is projected in the higher, external argument position). Pashto is therefore minimally different from Japanese, in which subjects and objects are equidistant to T due to verb-raising to T. It was noted in section 1.5.1 that there is no evidence that the verb moves overtly to T, and indeed the fact that both the subject and object are equidistant from T suggests that the verb does *not* move overtly to T. Evidence for this subject/object asymmetry with respect to the EPP is that scrambling of the object to the left of the subject induces reconstruction of the object NP. This is seen most clearly when an object anaphor, *khpel zaan* 'own self', is fronted. Consider the sentences in (116–117), all of which are grammatical. The (a) sentences illustrate SOV order, and the (b) sentences illustrate OSV order; the latter are of particular interest, as they show that the object anaphor

khpel zaan 'own self' reconstructs at LF in order to be locally A-bound by the subject, *spi/spay* 'dog'.

(116)   a.   spi   khpel   zaan   khog   kerr

dog(OBL MASC)   *own*   *self*   *hurt*   *do(PAST PERF MASC3SG)*

b.   khpel   zaan   spi   khog   kerr

own   *self*   *dog(OBL MASC)*   *hurt*   *do(PAST PERF MASC3SG)*

'the dog hurt himself' (Jan Mohammad, p.c.)

(117)   a.   spay   khpel   zaan   khog-aw-i

dog(DIR)   *own*   *self*   *hurt-TRANS(PRES IMPF)-3SG*

b.   khpel   zaan   spay   khog-aw-i

own   *self*   *dog(DIR)*   *hurt-TRANS(PRES IMPF)-3SG*

'the dog is hurting himself' (Jan Mohammad, p.c.)

The subject, being an external argument, is closer to Spec/TP than any other argument, and hence, by Shortest Move, must move there. Subsequent movement of the object to the left of the subject—as in the (b) sentences of (116–117) above—is therefore to a higher, A' position (presumably, as in Japanese, a focus projection).[42]

It should be stressed that the acceptability of the SOV and OSV word orders illustrated above are not because the anaphor is necessarily 'subject-oriented,' as has sometimes been claimed about the anaphor in Hindi. In Pashto, the anaphor may also take a non-subject (i.e.,

---

[42] Pashto's word order variants carry different focus interpretations (Tegey 1979: 379), but these have not been well studied. Penzl (1955: 133) and Tegey and Robson (1996: 176–177) identify a sentence-initial focus position, while Shafeev (1964: 55) and Babrakzai (1999: 63–64) report that focused elements immediately precede the predicate.

direct or indirect object) as its antecedent. In the following sentence, the full anaphor *khpel zaan* 'own self' prefers to take the nearer NP—the indirect object—as its antecedent (although the more distant subject remains available as an antecedent):

(118) parun   [S ásiya] [IO máryem ta] [DO **khpel zaan**]
       yesterday   Asia   Maryam   to   own   self

       po   burtshikhánee kee   wa   shayele
       in   kitchen   in   PERF   showed

       'Yesterday, Asia$_i$ showed herself$_{j/?i}$ to Maryam$_j$ in the kitchen' (Boraka)

Only by moving the reflexive DO to a position between the subject and IO is the preferred interpretation reversed:

(119) parun   [S ásiya] [DO **khpel zaan**] [IO máryem ta]
       yesterday   Asia   own   self   Maryam   to

       po   burtshikhánee kee   wa   shayele
       in   kitchen   in   PERF   showed

       'Yesterday, Asia$_i$ showed herself$_{i/*j}$ to Maryam$_j$ in the kitchen' (Boraka)

In another variety of Pashto, the possessive anaphor *khpel* 'self's' has similar behavior. Below, it may take either the subject or the indirect object as its antecedent:

(120) [S maa]   [IO Tor ta] [DO **khpel-a**   brakha]
       PN1SG(OBL)   Tor   to   self(POSS)-FEM DIR SG   share(FEM)

       war   krr-a
       3SG(DAT)   do(PAST PERF)-FEM3SG

       'I$_i$ gave Tor$_j$ self's$_{i/j}$ share' (Jan Mohammad, p.c.)

95

3.4.3. Distribution with verbal agreement suffixes

If clitics were arguments, they would be expected to co-occur with agreement suffixes on the verb. However, clitics and verbal agreement suffixes are in complementary distribution, suggesting that they have the same role: to bear features of (minimally) person and number, and license empty pronominals when they are topical.

Verbal agreement suffixes co-occur with null and overt direct-case arguments (nominative, absolutive), while clitics license null oblique-case arguments (ergative, accusative, genitive). The agreement suffixes on the verb are obligatory, and may license *pro*-drop of either the subject or direct object (depending on tense), or may co-occur with an overt NP argument. In present tense, the verb agrees with the subject (121a), while in past tense, the verb agrees with the object (121b).

(121)   a.   ahmad   ghwa   lwesh-**i**

               Ahmad(MASC)   *cow(FEM SG)*   *milk-3PRES*

               'Ahmad is milking the cow'

      b.   ahmad   ghwa   lwash-el-**a**

               Ahmad(MASC)   *cow(FEM SG)*   *milk-PAST-FEM3SG*

               'Ahmad was milking the cow' (Babrakzai 1999: 75)

The same obligatory agreement is triggered by strong pronouns. In (122), the strong 1sg pronoun *ze* co-occurs with the obligatory agreement suffix on the verb. Because this agreement suffix is rich enough to identify a 1sg subject, the pronoun *ze* may also be omitted.

(122)   (ze)   hara   wrez   pe   baagh   kee   gerz-**em**

          PN1SG   *every*   *day*   *at*   *garden*   *in*   *walk-1SG*

          'I walk in the garden every day' (Babrakzai 1999: 75)

Strong pronouns appear in the same positions as full NP arguments. Unlike clitic pronouns, they do not appear in second position. In (122), the strong 1sg pronoun *ze* appears at the beginning of the sentence—a position that a 2P clitic could never occupy, as it would lack a phonological host to its left.

With respect to agreement, clitics are unlike such full NPs and strong pronouns, which are arguments that are typically fixed with respect to the verb. Specifically, clitics are in complementary distribution with the obligatory agreement morphology on the verb. Clitics are underlined (as usual):

(123) a. gad-eed-**em**   (*<u>mee</u>)

       dance-INTR-1SG(PAST IMPF) *1SG*

       'I was dancing' (Yusufzai)

    b. khkol-ew-**i**   <u>mee</u> (*<u>yee</u>)

       kiss-TRANS-3SG(PRES IMPF) *1SG* *3SG*

       'he is kissing me' (Yusufzai)

    c. ahmad (*<u>mee</u>) khkol-ew-**em**

       Ahmad *1SG*   *kiss-TRANS-1SG(PAST IMPF)*

       'Ahmad was kissing me' (Yusufzai)

    d. ahmad (*<u>yee</u>) gad-íg-**i**

       Ahmad *3SG*   *dance-INTR-3SG(PRES IMPF)*

       'Ahmad is dancing' (Yusufzai)

If clitics were arguments, they would be expected to co-occur with agreement morphology, exactly as direct-case full NP arguments must, as in (121), and as strong pronouns must, as in (122). This same point is illustrated in (145) below, in which the 1sg clitic *mee* and the 1sg verbal agreement suffix *-em* must agree with distinct arguments.

This complementary distribution between verbal agreement suffixes and 2P pronominal clitics is evidence that both kinds of morphology serve to identify *pro*. This *pro* is homogenous in character, irrespective of whether it is identified by a verbal agreement suffix or by a pronominal clitic head. The kinds of asymmetries that obtain between subject and object *pro* in Chinese (Huang 1984), for example, do not obtain in Pashto. Consider the fairly free referential possibilities that are available to English (overt) pronouns:

(124)　　a.　He came

　　　　　b.　$Bill_i$ saw $him_{j/*i}$

　　　　　c.　$John_i$ said that $he_{i/j/*k}$ knew $Bill_k$

　　　　　d.　$John_i$ said that $Bill_j$ knew $him_{i/k/*j}$

The pronoun in (a) takes its referent from the discourse, as does the pronoun in (b)—the latter due to Condition B. When the pronoun appears in an embedded clause, the same facts obtain. Both the subject pronoun in (c) and the object pronoun in (d) may corefer with the matrix subject *John* or may take a distinct referent from discourse. There is thus no subject/object asymmetry with respect to the ability of an English pronoun to take a discourse referent. This is not the case with empty pronouns (*e*) in Chinese:

(125)　　a.　*e*　lai-le

　　　　　　　come-PERF

　　　　　'He came'

b. Lisi   hen   xihuan   *e*

   Lisi   very   like

   'Lisi$_i$ likes him$_{j/*i}$ very much'

c. Zhangsan   shuo   [*e*   bu   renshi   Lisi]

   Zhangsan   say   not   know   Lisi

   'Zhangsan$_i$ said that he$_{i/j/*k}$ did not know Lisi$_k$'

d. Zhangsan   shuo   [Lisi   bu   renshi   *e*]

   Zhangsan   say   Lisi   not   know

   'Zhangsan$_i$ said that Lisi$_j$ did not know him$_{k/*i}$'   (Huang 1984: 537)

As shown by the subscript indices in the translations, empty pronouns in Chinese have a similar distribution to strong pronouns in English, at least with respect to the (a) and (b) sentences. The (c) and (d) sentences, however, show that the empty pronoun in Chinese has a different distribution, depending on whether it is the subject or object. While the subject pronoun in the embedded clause of (125c) has the same distribution as its English counterpart in (124c), the object pronoun in (125d) has a more restricted distribution, being forbidden from taking the matrix subject as its antecedent; rather, it must take a distinct referent from discourse. In order for the object in (125d) to refer to the matrix subject or to someone else, a strong pronoun is used:

(126)   Zhangsan   shuo   [Lisi   bu   renshi   ta]

        Zhangsan   say   Lisi   not   know   him

        'Zhangsan$_i$ said that Lisi$_j$ did not know him$_{i/k}$'   (Huang 1984: 538)

Because the empty object pronoun in Chinese cannot be A-bound by a matrix argument, but may be A'-bound by a topic in the discourse, Huang (1984) suggests that the null object in Chinese is a variable (not a pronoun), which is A'-bound by a null topic.[43]

Such asymmetries do not obtain in Pashto among the empty categories assumed to be identified by verbal agreement suffixes and clitics. Each of the sentences in (127) has an object represented differently in the embedded clause: the 3sg clitic *yee* in (a), the feminine 3sg strong pronoun *daa* in (b), and a null object identified by 3sg feminine verbal agreement (c). As indicated by the subscript indices, the object clitic in (a) and the null object in (c) have the same referential possibilities, while the strong pronoun in (b) has a more restricted distribution:

(127) a. mina   peekar   kaw-i   [tshee Tor   <u>ba</u>   yee   bóz-i]

           Meena   thought do-3SG COMP Tor   FUT   3SG take(PERF)-3SG

           'Meena$_i$ thinks that Tor will take her$_i$'

      b. mina   peekar   kaw-i   [tshee Tor   <u>ba</u>   **daa**   bóz-i]

           Meena   thought do-3SG COMP Tor   FUT   PN3SG(FEM)   take-3SG

           'Meena$_i$ thinks that Tor will take her$_{j/*i}$'

      c. mina   peekar   kaw-i   [tshee Tor   bótl-**a**]

           Meena   thought do-3SG COMP Tor   took(PERF)-FEM3SG

           'Meena$_i$ thinks that Tor took her$_i$' (Jan Mohammad, p.c.)

Using a pronominal clitic or a null object, as in (a) and (c) respectively, is the unmarked way of expressing such sentences (Jan Mohammad, p.c.); the only factor determining the choice between a pronominal clitic or a null object is ergativity: in non-past tense (a), the verbal agreement

---

[43] This asymmetry between empty categories in subject vs. object position is also observed in Korean and Brazilian Portuguese (Huang 1984: 540–541).

identifies the subject, and so the clitic represents the object; in past-tense (c), the verbal agreement identifies the object, which may be omitted. Using a strong pronoun in (b) introduces a focused referent that must be disjoint from the matrix subject.

The same asymmetry obtains between null and overt pronouns in subject position:

(128)   a.   Tor   peekar   kaw-i   [tshee   sabaa   <u>ba</u>   laarr   sh-**i**]

Tor   *thought*   *do-3SG*   *COMP*   *tomorrow*   *FUT*   *go*   *become-3SG*

'Tor$_i$ thinks he$_i$ will go tomorrow'

b.   Tor   peekar   kaw-i   [tshee   **hagha**   <u>ba</u>   sabaa

Tor   *thought*   *do-3SG*   *COMP*   *PN3SG*   *FUT*   *tomorrow*

laarr   sh-i]

go(PERF)   *become(PERF)-3SG*

'Tor$_i$ thinks he$_{j/*i}$ will go tomorrow' (Jan Mohammad, p.c.)

The 3sg verbal agreement suffix *-i* identifies the subject in both sentences. This agreement licenses subject *pro*-drop in (a), in which the dominant interpretation for the omitted NP is for it to corefer with the matrix subject. If a strong subject pronoun appears in the embedded clause, as in (b), the pronoun is disjoint in reference from the matrix subject: again, because the strong pronoun brings a distinct referent into focus. This asymmetry is familiar from better-studied *pro*-drop languages like Spanish:

(129)   a.   Gabi$_i$ piensa   que   *pro$_i$*   vio   a   Pablo

Gabi   *thinks*   *COMP*   *saw(3SG)*   *ACC*   *Pablo*

b.   Gabi$_i$ piensa   que   ella$_{j/*i}$ vio   a   Pablo

Gabi   *thinks*   *COMP*   *PN3SG(FEM)*   *saw(3SG)*   *ACC*   *Pablo*

'Gabi thinks she saw Pablo' (María Cristina Cuervo, p.c.)

With respect to the ability of a pronominal embedded subject to refer to the matrix subject, Pashto and Spanish therefore accord the same distributions to strong versus weak (*pro* and clitic) pronouns.

For Huang (1984), forbidding *pro* in object position allows Chinese and Romance languages to be grouped together in forbidding object *pro*, since the empty object in Chinese is argued not to be *pro*, but rather a variable—thereby allowing Huang to ask of this typology of languages: '... why is it that all the languages under investigation allow only zero subject pronouns, but exclude zero object pronouns?' (546). In posing this question, however, he is also explicit in excluding Pashto from further consideration: 'Henceforth, I will concentrate on languages with no agreement [Chinese] or only subject-verb agreement [Romance], and will not consider languages like Pashto. It should be remembered that when I claim that an object EC cannot be a pronominal, I do not refer to languages showing verb-object agreement' (Huang 1984: 545 n. 13). The Pashto asymmetries in (127–128) provide further evidence that clitics and verbal agreement have a similar function, which is to license (identify) the same empty pronominal (*pro*). The choice between verbal agreement and clitic agreement is determined simply by tense: in non-past tense, the verb agrees with the subject, and so a clitic is used to identify the object; in past tense, the verb agrees with the object, and so a clitic is used to identify the subject. That both verbal agreement and clitics have the same referential possibilities, as shown above, groups them in opposition to strong pronouns, which not only appear in fixed argument positions, but have different discourse interpretations.

This analysis accords perfectly with Huang's (1984: 535–536) own description of Pashto, in which verbal agreement licenses both subject- and object-*pro*-drop, as determined by tense, without the subject-object asymmetries that occur in Chinese. The data of (127–128) take the point one step further by showing that there are likewise no referential asymmetries between clitics, on the one hand, and *pro* when it is identified by verbal agreement; this empty category is

homogeneously pronominal, rather being an A'-bound variable.[44] It is simplest, then, to group clitics and verbal agreement together in their syntactic function, which is to license *pro*. Distinguishing among these empty categories is not necessary for Pashto as it is for Chinese.[45]

3.4.4. Coordination

Asymmetries in coordination also suggest that clitics should not be treated as arguments. As illustrated below, *aw* 'and' may conjoin full NPs and strong pronouns with themselves or with each other:

(130) a. parun aman aw asad kitab olwelewu

           yesterday *Aman* *and* *Asad* *book* *read*

           'Aman and Asad read a book yesterday' (Yusufzai)

      b. parun taa aw maa kitab olwelewu

           yesterday *PN2SG* *and* *PN1SG* *book* *read*

           'you and I read a book yesterday' (Yusufzai)

      c. parun Aman aw maa kitab olwelewu

           yesterday *Aman* *and* *PN1SG* *book* *read*

           'Aman and I read a book yesterday' (Yusufzai)

---

[44] Sec. 3.6.2 will show that *pro* may be resumptive, being A'-bound by a null operator in relative clauses and left-dislocation structures.

[45] This conclusion should not be surprising, as null objects in Hindi—a language closely related to Pashto—also are pronominal, rather than variables bound by a null operator (Dwivedi 1994: 43–46). See sec. 3.6.2 below for evidence from relative clauses and left-dislocation that the empty category identified by a clitic is *pro*, rather than trace.

       d.   parun   maa   aw   Aman   kitab   olwelewu

             yesterday   *PN1SG*   *and*   *Aman*   *book*   *read*

            'Aman and I read a book yesterday' (Yusufzai)

In contrast, clitic pronouns may not be conjoined. The following sentences each have an individual pronominal clitic in second position:

(131)   a.   parun   <u>dee</u>   kitab   olwelewu

             yesterday   *2SG*   *book*   *read*

            'you read a book yesterday' (Yusufzai)

      b.   parun   <u>mee</u>   kitab   olwelewu

             yesterday   *1SG*   *book*   *read*

            'I read a book yesterday' (Yusufzai)

The clitic pronouns may not be conjoined with each other, however:[46]

(132)   a.   *parun   <u>dee</u>   aw   <u>mee</u>   kitab   olwelewu

             yesterday   *2SG*   *and*   *1SG*   *book*   *read*

            'you and I read a book yesterday' (Yusufzai)

      b.   *parun   <u>mee</u>   aw   <u>dee</u>   kitab   olwelewu

             yesterday   *1SG*   *and*   *2SG*   *book*   *read*

            'you and I read a book yesterday' (Yusufzai)

---

[46] See Tegey (1977: 19) for other examples.

Nor may a clitic pronoun be conjoined with a full NP; the following sentences thus contrast with (130c–d):

(133) a. *parun Aman aw <u>mee</u> kitab olwelewu

     yesterday *Aman* *and* *1SG* *book* *read*

    'Aman and I read a book yesterday' (Yusufzai)

    b. *parun <u>mee</u> aw Aman kitab olwelewu

     yesterday *1SG* *and* *Aman* *book* *read*

    'Aman and I read a book yesterday' (Yusufzai)

If clitics represented arguments, they should be able to be conjoined, as may full NPs and strong pronouns. It is not plausible that it might be merely a phonological quirk of clitics that they cannot be conjoined, since the sentences in (112–113) above have already demonstrated that clitics in embedded clauses may be hosted by a functional category as light as the complementizer *tshee*, if the syntax derives such a structure. These asymmetries from coordination, then, suggest again that clitics do not behave as regular arguments. If clitics are treated as agreement morphemes, however, there would be no more reason to expect them to be able to be conjoined than if they were more 'canonical' agreement morphemes such as appear as suffixes on the verb.

3.4.5. Doubling

The phenomenon of clitic doubling—which has long been taken in analyses of Spanish clitics as the best evidence that clitics represent agreement—also exists in Pashto with a separate set of verbal clitics. Whereas 2P pronominal clitics correspond to full NPs that would be marked ergative, accusative, and genitive, verbal pronominal clitics correspond to NPs that would be marked dative, or as other sorts of indirect arguments; examples of such verbs are *leegel* 'send', *khayel* 'show', and *bakhel* 'give (as a gift)' (Babrakzai 1999: 82). Whereas the 2P pronominal

clitics distinguish person and number, the verbal clitics distinguish only person (Tegey 1977: 7, 34, 222; Tegey and Robson 1996: 65; Babrakzai 1999: 33, 81):

(134)   Verbal clitics

        raa   1SG/PL

        der   2SG/PL

        wer   3SG/PL

The verbal clitics have in common with the 2P clitics the property that they may not bear stress or be focused, in contrast to their corresponding strong forms. However, several properties of verbal clitics distinguish them from the 2P clitics: (i) they are proclitic to the verb; (ii) they are insensitive to tense (i.e., past-tense ergativity is irrelevant to the appearance of verbal clitics in a sentence); and (iii) they may double an overt NP. The following examples illustrate these properties:

(135)   a.   ahmad   (taa   ta)   yew   kitaab   **der**   kaw-i

           Ahmad   *PN2SG*   *to*   *one*   *book*   *2SG(DAT)*   *do-3SG*

           'Ahmad is giving you a book'

      b.   minee   (mug   ta)   meewa   **raa**   we   leegel-a

           Meena   *PN1PL*   *to*   *fruit*   *1PL(DAT)*   *PERF*   *sent-FEM3SG*

           'Meena sent us fruit' (Babrakzai 1999: 82)

As indicated by the parentheses, the indirect object PP, assumed to be the actual argument of the verb, may be omitted. Just as verbal agreement suffixes license optional *pro*-drop of nominative/absolutive arguments, the verbal clitics license optional omission of the indirect

object. Note also that the clitic *raa* in (135b) precedes the perfective morpheme *we*, which is entirely expected if dative clitics are generated in a VP-external agreement projection.

Very little will be said here about verbal clitics, as they are a topic unto themselves (see Tegey 1977: ch. 7). What is directly relevant to the present analysis of 2P clitics is the following asymmetry:

(136)   Verbal agreement suffixes: optional doubling of NOM/ABS

Verbal clitics: optional doubling of DAT

Second-position clitics: complementary distribution with ERG/ACC/GEN

If 2P clitics represent properties of agreement, we should expect them to double full NPs, as do the more canonical agreement morphemes. It was noted in chapter 1 that the co-occurrence of full NPs with verbal agreement suffixes is expected if verbs entered the derivation as fully formed words, and hence do not have the option to lose or be separated from their inflectional morphology. Verbal clitics differ from agreement suffixes only in this property: they are merged as autonomous agreement heads, and hence are expected to host full NPs in their specifiers. The problem is therefore to explain why 2P clitics may not likewise host full NPs in their specifiers, but rather are in complementary distribution with full NPs.

The least desirable solution to this problem would be to assume that 2P clitics are arguments, given the evidence against this possibility that has so far been presented. A possibly related asymmetry is found in Romance languages. In Spanish, clitic doubling is always permitted with dative arguments, and is obligatory with pronominal datives:

(137)   a.   Miguelito   **le**   regaló   un   caramelo   (**a   Mafalda**)

Miguelito   3SG(DAT)   gave   DET   candy   to   Mafalda

'Miguelito gave Mafalda a (piece of) candy'

b. **Le** entregué la carta **a él**

   3SG(DAT) *delivered* *DET* *letter* *to* *3SG.MASC.OBL*

   'I delivered the letter to him'

c. *Entregué la carta **a él**

   delivered *DET* *letter* *to* *3SG.MASC.OBL*

<div align="right">(Jaeggli 1982: 12–13)</div>

Aside from the pronominal/non-pronominal asymmetry, the behavior of datives in Spanish is therefore identical to their behavior in Pashto. So too is the behavior of accusative NPs in standard varieties of Spanish, French, and Italian similar to Pashto. As the following Spanish sentences show, full NP accusatives may not co-occur with accusative clitics:

(138)   a.   Vimos la casa de Mafalda

        saw(1PL) *DET* *house* *POSS* *Mafalda*

        'we saw Mafalda's house'

        b.   ***La** vimos la casa de Mafalda

            3SG(ACC FEM) *saw(1PL)* *DET* *house* *POSS* *Mafalda*

(139)   a.   Vimos a Guille

        saw(1PL) *to* *Guille*

        'we saw Guille'

        b.   ***Lo** vimos a Guille

            3SG(ACC MASC) *saw(1PL)* *to* *Guille*

(140)   a.   La vimos

                     'we saw her/it'

             b.   Lo vimos

                     'we saw him/it'

<div align="right">(Jaeggli 1982: 14)</div>

In River Plate Spanish, however, clitic doubling is preferred when the accusative is animate and specific. The sentence in (139b) above is therefore grammatical in that variety of Spanish. In all varieties of Spanish, pronominal accusative NPs are obligatorily doubled by a clitic:

(141)   a.   Lo   vi   a   él

                     3SG(ACC MASC)   *saw(1SG)*   *to*   *3SG.MASC.OBL*

                     'I saw him'

             b.   *vi   a   él

                     saw(1SG)   *to*   *3SG.MASC.OBL*

<div align="right">(Jaeggli 1982: 14)</div>

These data show that an object NP may be doubled by a clitic only if the NP is preceded by a preposition—what Jaeggli (1982: 20) calls 'Kayne's Generalization'—and indeed this observation suggests a uniform account of direct and indirect objects in Spanish. Jaeggli (1982: 22) suggests that clitics absorb the case assigned by the verb, and so any overt NPs that are doubled—whether they are accusative or dative—must be preceded by a preposition in order to receive case themselves.

While it is desirable to avoid the idea that clitics absorb case (because of its implication that clitics are argument-like), a variant of Jaeggli's analysis may be applied straightforwardly to

<div align="center">109</div>

datives in Pashto. When an overt dative argument appears, it is invariably followed by the case-assigning postposition *ta* 'to', as illustrated above in (135), and so clitic doubling is permitted in such constructions; the clitic is merely another VP-external agreement projection, but one that is lower than the oblique agreement projection that checks the person and number features of direct arguments (ergative, accusative) and genitives. When clitic-doubling does not occur, *pro* is therefore caseless, but because it is not an overt NP, it escapes whatever gives the effect of the Case Filter (which only holds of overt NPs). The *pro* argument identified by the 2P ergative, accusative, and genitive clitics, on the other hand, does bear case, as may be seen from the ability of a corresponding overt NP to appear alone without an agreeing clitic or a postposition. But the complementarity between overt direct case (ergative, accusative, genitive) arguments and clitics remains an intriguing area for more detailed study.

3.5. Possessive clitic dislocation

The genitive function of second-position clitics provides further evidence that clitics are agreement morphemes, as well as offering broader insight into the organization of grammar that derives the interpretation of clitics at LF. By constructing sentences having different combinations of overt NPs and 2P clitics, the clitics may be compelled to have either fixed or free interpretations. Schematically, some relevant configurations are the following, where NP represents a phonetically overt nominal (i.e., not *pro*):

(142) a. NP_i cl $V_{[trans]i}$

     b. NP_i cl cl $V_{[trans]i}$

     c. NP_i cl NP $V_{[trans]i}$

     d. NP_i cl cl NP $V_{[trans]i}$

All of these configurations contain a transitive verb, which therefore requires two arguments. The verb in each case agrees with the overt NP, as indicated by the subscript index; i.e., none of these

cases involve *pro*-drop. In (a), then, the single NP occupies one argument position, compelling the clitic to occupy the other. In (b), a second clitic is added; in such cases, ambiguous interpretations arise, because either clitic may serve as the second argument or as the possessor of the overt NP; this configuration reveals that clitics are not ordered according to their case or function, as they are in more familiar languages. In (c), there are two overt NPs, and so the argument structure of the verb is saturated; in this configuration, the single clitic is obliged to be genitive, taking the highest NP as its possessum. Despite the ambiguity in (b), (c) shows that genitive clitic interpretation obeys a simple locality constraint: it associates with the highest overt NP in its clause. The (a–c) configurations are examined in this section. Finally, in (d), the verb's argument structure is saturated by two overt NPs, and so both of the clitics are forced into genitive functions; unlike in (b), there is no ambiguity of clitic interpretation, and each clitic associates with its nearest NP. This last case will be explored in the following chapter, where it will be taken as evidence that second-position clitics are merged directly into their surface positions, rather than being ordered by a morphological template after Spell Out.

This section is sub-divided as follows. Section 3.5.1. presents evidence from the behavior of genitive clitics that continues to suggest that the clitics are agreement morphemes, rather than arguments. Section 3.5.2. compares superficially similar constructions in other languages, explaining why Pashto's clitics warrant a novel analysis. Section 3.5.3. shows how the present analysis—in which *pro* moves to the specifier of an agreement projection (headed by the clitic) in order to erase uninterpretable features—accounts for both the interpretation and placement of the clitics. Although the empty pronominal can appear to move rather freely, section 3.5.4. shows that its movement is in fact local (and hence its interpretation correspondingly fixed). The net effect is, again, that if second-position clitics are analyzed as agreement morphemes, their placement and interpretation is explained without recourse either to syntactic movement of the clitics themselves or to phonological requirements concerning their placement. Second-position clitics may be merged directly into their surface position, heading structurally high (and, in

principle, freely generated) agreement projections. This analysis results in a simple account of why second-position clitics appear at all, and where they are positioned in the sentence.

3.5.1. More evidence for clitics as agreement

The most obvious property of pronominal clitics that makes them look distinctly unlike arguments (in comparison to full NPs and strong pronouns) is that they do not appear in the positions in which arguments typically appear. Recall from (11) in chapter 1 that strong possessive NPs are rigidly positioned immediately before the heads they modify. In contrast, possessive clitics need not be adjacent to the nouns they modify. Consider the 1sg possessive clitic *mee* in the phrase 'my father' in the following sentences. In (143), the possessive clitic follows the head noun:

(143)   plaar   <u>mee</u>   byáy-em

    father   *1SG*   *bring-1SG(PRES)*

    'I am bringing my father' (Tegey and Robson 1996: 174)

When the sentence has another NP in initial position, however, the clitic appears to the left of its possessum, taking instead the initial NP as its phonological host:

(144)   khoshal   khan   <u>mee</u>   plaar   day

    Khoshal   Khan   1SG   father   be(MASC3SG)

    'Khoshal Khan is my father' (Tegey and Robson 1996: 174)

When the possessum is embedded inside a PP, as it is in (145), the clitic appears entirely removed from the NP, taking the PP (rather than the NP with which it is semantically associated) as its phonological host:

(145)   [_{PP} le   plaar   sara]   <u>mee</u>   naast   y-em

          with   father   with   1SG   sitting(MASC SG)   be-1SG

          'I am sitting with my father' (Tegey and Robson 1996: 174)

The contrast between (143) and (145) is especially revealing: if the 1sg possessive clitic *mee* may take *plaar* 'father' as its phonological host in (143), it should also be expected to do so in (145)—especially considering that in (145), it is exactly *plaar* 'father' with which the 1sg possessive clitic *mee* is semantically associated, and that this NP appears initially in the clause (within the PP). Indeed, ambiguity can arise, exactly because the base position of a clitic may be obscured by the fixed order of clitics with respect to each other inside the 2P clitic cluster, which was schematized above in (86).[47] In this respect, consider the ambiguity of the following sentence:

(146)   plaar   <u>mee</u>   <u>dee</u>   léeg-i

          father   1SG   2SG   send-PRES3SG

          'My father is sending you' or 'Your father is sending me' (Tegey and Robson 1996: 175)

If context does not resolve such ambiguities, strong pronouns must be used, since they appear in fixed positions (Tegey and Robson 1996: 176). The following sentence is unambiguous:

(147)   [di   taa   plaar]   <u>mee</u>   <u>dee</u>   wé-leg-i

          poss   PN2SG   father   1SG   must   PERF-send-PRES3SG

          'Your father must send me' (Tegey and Robson 1996: 176)

---

[47] Bubenik (1994: 113, 119–120) discusses similar examples in Middle Persian. Such ambiguities are rare, but are not confined to Iranian; subject/object enclitics in Southern Paiute are homophonous with the possessive forms, creating similar ambiguities (Sapir 1930: 187; 1992: 205).

Note that the clitic *dee*, which normally may have either a 2sg or a modal ('must') interpretation, must be interpreted as a modal in (147), since the transitive verb only projects two arguments, and 2P clitic pronouns may not corefer with full NPs or strong pronouns (i.e., there is no clitic doubling).

Because of their reasonably fixed position with respect to the verb, it is simplest to assume that full NPs and strong pronouns are merged in argument positions, and move only in order to check EPP or topic/focus features. The default word order of sentences having full NPs and strong pronouns is SOV, and certainly in any sentence in which the verb is not focused, full NPs and strong pronouns never follow the verb (cf. Tegey 1977: 238 ff.). In this respect, full NPs and strong pronouns have a fixed order with respect to the verb: at the very least, they must precede it. Clitics differ in this regard: they do not have a fixed position with respect to the verb, but rather only require a phonological host, of any syntactic category, to their left. The possessive clitic therefore appears to be removed from its semantic domain (NP/DP), appearing in a larger syntactic domain (TP/IP). This property will be called 'possessor dislocation,' in order to distinguish it from more familiar 'possessor raising' and 'possessive dative constructions,' which will be seen to have entirely different properties from those found in Pashto. The broader term 'external possession' will also be avoided here, due to its similar implication that the displaced possessor is an argument of the verb: 'We take core instances of external possession to be constructions in which a semantic possessor-possessum relation is expressed by coding the possessor as a core grammatical relation of the verb and in a constituent separate from that which contains the possessum' (Payne and Barshi 1999b: 3).

Another kind of possessor dislocation is observed in Modern Irish (McCloskey and Hale 1983). The language is rigidly head-initial, and so genitive full NPs follow their possessees:

(148)   a.   bád   beag   Eoghain

         boat   little   Owen(GEN)

         'Owen's little boat'

         b.   teach   beag   suarach   thuismitheoirí   Eoghain

         house   little   wretched   parents   Owen(GEN)

         'Owen's parents' wretched little house' (McCloskey and Hale 1983: 511)

Oddly, however, pronominal possessors precede the head noun:

(149)   a.   mo   theach

         my   house

         b.   bhur   dteach

         your(PL)   house

McCloskey and Hale (1983) suggest that this asymmetry is only apparent: in phrases like (149), there is argued to be a post-nominal possessor *pro* (in the canonical, post-head argument position), while the prenominal possessor is taken to represent agreement morphology, analogous to the person-number agreement that also appears on verbs and prepositions, as schematized below for (149a):

(150)

```
                    NP
              ┌──────┴──────┐
              N             NP
          ┌───┴───┐         │
         Agr      N        PRO
          │       │        1SG
          │       │
         mo     teach
         1SG    house
```
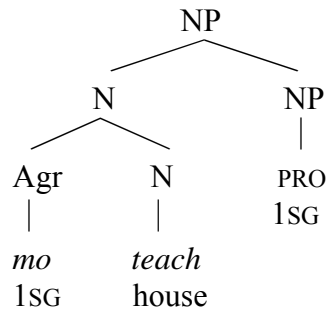
(adapted from McCloskey and Hale 1983: 513)

As McCloskey and Hale (1983: 525) note, this sort of analysis has the result of eliminating the theoretical difference between agreement and cliticization, an idea that has been advanced more recently by Sportiche (1996). Such an analysis is also well suited for Pashto's second-position clitics, with a notable difference being that all agreement projections identifying *pro* appear within TP, whether *pro* is a subject, object, or possessor. As explained above, this analysis allows the syntax to derive the bulk of Pashto's second-position effects without the need to postulate rules of clitic movement.

If Pashto's possessive clitic were treated as an argument, it would be difficult to explain why its movement from its base position (Spec/NP or Spec/DP) into the clause is grammatical, given that overt possessor extraction is independently excluded in the language. Extracting a possessive pronominal violates the Left Branch Condition (151), which excludes (among other things) such English sentences as those in (152):

(151)   Left Branch Condition

No NP which is the leftmost constituent of a larger NP can be reordered out of this NP by a transformational rule. (Ross 1967: 127)

(152)   a.   *Whose$_i$ do you like t$_i$ mother?

b.   *Which$_i$ did you read t$_i$ book?

There are nevertheless languages that permit overt violations of the Left Branch Condition, such as Russian (153–154) and Latin (155):

(153)   a.   C,ju   knigu   ty   citaješ?
             whose   book   you   read
             'Whose book are you reading?'

        b.   **C,ju**   ty   citaješ   **knigu**?
             whose   you   read   book
             'Whose book are you reading?' (Ross 1967: 145)

(154)   a.   Skol,ko   let   emu   bylo?
             how many   years   to-him   be
             'How many years old was he?'

        b.   **Skol,ko**   emu   bylo   **let**?
             how many   to-him   be   years
             'How many years old was he?' (Ross 1967: 145)

(155)   **Cuius**   legis   **librum**?
        whose   read-you   book
        'Whose book are you reading?' (Ross 1967: 145)

Ross (1967: 145–146) notes that highly inflected languages with scrambling are among the most likely to permit Left Branch Constraint violations. Other languages that permit possessor extraction include Hindi (156–157) and Slovenian (158). The *wh*-questions in which the possessor alone is fronted are grammatical with an echo interpretation:

117

(156) a. tum-ne   kis-kii   kitaab   paRh-ii   thii

           you-ERG   *who-GEN.F* *book.F*   *read-PERF.F* *PAST*

           'Whose book had you read?'

       b. **kis-kii**,   tum-ne   **kitaab**   paRh-ii   thii

           who-GEN.F   *you-ERG*   *book.F*   *read-PERF.F* *PAST*

           'Whose book had you read?' (Rajesh Bhatt, p.c.)

(157) a. tum-ne   Ram-kii   kitaab   paRh-ii   thii

           you-ERG   *Ram-GEN.F* *book.F*   *read-PERF.F* *PAST*

           'You had read Ram's book'

       b. **Ram-kii**,   tum-ne   **kitaab**   paRh-ii   thii

           Ram-GEN.F   *you-ERG*   *book.F*   *read-PERF.F* *PAST*

           'As for Ram, you had read his book' (Rajesh Bhatt, p.c.)

(158) a. C"igavo   knjigo   si   prebral?

           whose   *book*   *be(PRES2SG)*   *read*

           'Whose book did you read?'

       b. **C"igavo**   si   prebral   **knjigo**?

           whose   *be(PRES2SG)*   *read*   *book*

           'Whose book did you read?' (Tatjana Marvin, p.c.)

Although Pashto is like these languages in having rich inflection and scrambling, overt possessor

extraction is disallowed in Pashto: phonologically overt possessors may not be separated from

their possessees, as is shown in (159) for a possessive *wh*-word—contrasting strikingly with (156) from the closely related language Hindi.[48]

(159) a. taa   **de**   **tsha**   **kitab**

PN2SG(OBL)   POSS   who(OBL)   book

pe   ashpazkhana   kee   we   lwest

in   kitchen   in   PERF   read(PAST3SG)

'whose book did you read in the kitchen?'

b. **de**   **tsha**   **kitab**   taa

POSS   who(OBL)   book   PN2SG(OBL)

pe   ashpazkhana   kee   we   lwest

in   kitchen   in   PERF   read(PAST3SG)

'whose book did you read in the kitchen?' (Laghman)

c. ***de**   **tsha**   taa   **kitab**

POSS   who(OBL)   PN2SG(OBL)   book

pe   ashpazkhana   kee   we   lwest

in   kitchen   in   PERF   read(PAST3SG)

At best, extracting a quantifier is marginally grammatical:

(160) a. áhmad   **tso**   **kitaab-úna**   lar-i

Ahmad   how.many   book-PL   have-PRES3SG

'How many books does Ahmad have?'

---

[48] Thanks to Jan Mohammad for trying, without success, to construct a grammatical Left Branch Violation in Pashto.

b.   ?**tso**   áhmad   **kitaab-úna**   lar-i[49]

        how.many   *Ahmad*   *book-PL*   *have-PRES3SG*

        'How many books does Ahmad have?' (Farooq Babrakzai, p.c.)

(161)   a.   taa   **tsomra**   **ghanem**   wáa-khist-el

              2SG(ERG)   *how.much*   *wheat(MASC PL)*   *PERF-buy-PAST(MASC PL)*

              'How much wheat did you buy?'

        b.   ??**tsomra**   taa   **ghanem**   wáa-khist-el

              how.much   *2SG(ERG)*   *wheat(MASC PL)*   *PERF-buy-PAST(MASC PL)*

              'How much wheat did you buy?'

        c.   ?**tsomra**   <u>dee</u>   **ghanem**   wáa-khist-el[50]

              how.much   *2SG(ERG)*   *wheat(MASC PL)*   *PERF-buy-PAST(MASC PL)*

              'How much wheat did you buy?' (Farooq Babrakzai, p.c.)

Violations of the Left Branch Condition in overt syntax are therefore excluded; they are permitted only with the empty category *pro* (i.e., the possessive argument identified by a genitive clitic). If the possessive clitic is an agreement morpheme, and the actual possessor is the empty category *pro*, violations of the Left Branch Condition need not arise in overt syntax, as there is no need for overt movement of the possessor. Because the possessor is phonologically null, it is impossible to demonstrate from word order effects whether it moves in overt syntax—but its movement is

---

[49] Farooq Babrakzai (p.c.) describes this sentence as 'borderline grammatical'.

[50] This (c) sentence is more natural than the (b) sentence (Farooq Babrakzai, p.c.), suggesting again that clitics and strong pronouns indeed occupy different syntactic positions. The improved status of (c) over (b) is probably due to prosodic inversion of the clitic (see ch. 4).

surely covert, since cyclicity is independently known not to apply at LF, as was noted above with respect to (103). By Procrastinate, then, the possessor defers movement until LF, a level at which, crucially, (i) the necessary interpretation for *pro* and the pronominal clitics must occur anyway; and (ii) cyclicity does not apply. The only overt movement of *pro* is when it is nominative or absolutive (i.e., identified by a verbal agreement suffix), in which cases it moves to Spec/TP in order to satisfy the EPP (in the absence of an overt object).

Before turning to the details of this analysis, the following section examines external possession in some more familiar languages, showing that their properties are not comparable to those found in Pashto.

3.5.2. External possession in other languages

The dislocation of the possessive clitic from its possessee has the appearance of what has been called possessor raising or external possession in other languages: the possessor does not appear in its expected position or form, but rather is separated from its possessee, or appears in a different form.[51] A comparison between possessive 2P clitics in Pashto and possessor raising constructions in other languages reveals that Pashto's possessor dislocation construction is not comparable. This conclusion should not be surprising, as the evidence has been mounting that 2P clitics are agreement morphemes rather than arguments, and in many possessor-raising languages, possessive full NPs (arguments) may be removed from their possessums.

Kurdish, an Iranian language like Pashto, attaches its genitive clitics directly to the relevant NP. In some cases, however, 'genitive stranding' may also occur in Kurdish—though its behavior differs from that in Pashto, since the genitive clitic is also separated from the 2P clitics (which have the same form as the genitive clitics, as in Pashto, but do not serve genitive functions when they are in second position). Kurdish also differs from Pashto is permitting genitive stranding with some predicates, but not with others. For this reason, VanLoon (1997: 169), in passing, likens Kurdish's genitive stranding to possessive dative constructions in French. Because

---

[51] See Payne and Barshi (1999a) for a sampling of similar constructions in other languages.

Pashto's possessive dislocation might also appear, at first glance, to resemble possessive dative constructions, the following section compares these constructions in other languages, showing that their similarity to Pashto's possessive dislocation is only superficial. Pashto thereby merits a novel analysis.

3.5.2.1. Possessive dative constructions

The following sentences exemplify possessor raising in more familiar languages. Alongside the genitive constructions in the (a) sentences below, there is a corresponding possessive dative construction (b), in which the possessor appears as a dative argument of the verb, rather than as a genitive associated with its possessee:

(162)   Hebrew

    a.   Gil   higdil   et   ha-tmuna   **šel**   **Rina**

        Gil   *enlarged*   *ACC*   *the-picture*   *of*   *Rina*

        'Gil enlarged Rina's picture'

    b.   Gil   higdil   **le-Rina**   et   ha-tmuna

        Gil   *enlarged*   *to-Rina*   *ACC*   *the-picture*

        'Gil enlarged Rina's picture' (Landau 1999: 5)

(163)   French

    a.   J'ai   coupé   les   cheveux   **de**   **Pierre**

        I   *cut*   *the*   *hair*   *of*   *Pierre*

        'I cut Pierre's hair'

b. J'ai coupé les cheveux **à** **Pierre**

I *cut* *the* *hair* *to* *Pierre*

'I cut Pierre's hair' (Landau 1999: 3)

(164) Spanish

a. Revisé los informes **de** **los** **estudiantes**

I-revised *the* *reports* *of* *the* *students*

'I revised the students' reports' (María Cristina Cuervo, p.c.)

b. Les revisé los informes **a** **los** **estudiantes**

to-them *I-revised* *the* *reports* *to* *the* *students*

'I revised the students' reports' (Landau 1999: 3)

The interest posed by these variants has been stated thus (Landau 1999: 2):

(165) The classical puzzle of possessive datives

An argument in the clause (the possessor) derives its *semantic* role from another argument (the possessee), but its *syntactic* behavior from the predicate. What is the possessive dative an argument of?

The puzzle of Pashto possessive clitics is not the same. The question posed at the end of (165) is not applicable to the present analysis of Pashto, in which 2P clitics are agreement morphemes, not arguments. It is exactly this point that highlights the difference between the familiar possessor raising constructions and Pashto's possessive clitics. In all of (162–164), the genitive construction and the corresponding possessive dative construction have the same, full NPs as possessors, making it feasible to entertain the idea that the possessive datives are arguments of the

123

verb. Pairs like (162–164) do not exist in Pashto, since full NP possessors invariably appear in their base positions (preceding their possessums), as complements of the preposition *dee* 'of':

(166)   dee   asad   dee   plaar   tsalor   ddeeree   khaysta   lunnee

                 POSS   *Asad*   *POSS*   *father*   *four*   *very*   *pretty*   *daughters*

                 'Asad's father's four very pretty daughters' (Tegey and Robson 1996: 171)

Possessive clitics, on the other hand, always appear in second position, and furthermore do not have corresponding strong forms that appear in second position. Rather, the corresponding strong pronouns appear in their base positions, exactly as do full NPs, as was discussed above with respect to (146–147). The most salient difference between Pashto and the other languages, then, is that full NP and strong pronoun possessors appear in their expected positions, preceding their possessees. The possessors that appear in second position are not full NPs or strong pronouns, but are drawn only from the small set of highly underspecified clitic pronouns, which bear only person and number features.

There are further differences between Pashto and the other languages pertaining to possessor raising, which emphasize the point that the relevant constructions are not comparable. Landau (1999) observes that in possessive dative constructions, any VP-internal argument may serve as the possessee, whereas external arguments may not—entailing that a verb's compatibility with the possessive dative construction is a fairly reliable test of unaccusativity. In the following Hebrew sentences, the verb 'disappear' takes a VP-internal (unaccusative) argument—hence it is compatible with the possessive dative construction—whereas the verb 'run around' takes a VP-external (unergative) argument, making it incompatible with the possessive dative construction:

(167)   a.   ha-kelev   ne'elam   le-Rina

                 the-dog   *disappeared*   *to-Rina*

                 'Rina's dog disappeared'

b. *ha-kelev   hitrocec   le-Rina

   the-dog   ran-around   to-Rina

   'Rina's dog ran around' (Landau 1999: 7)

Nor may a possessive dative associate with relational nouns like 'brother' or 'mother' unless a genitive pronoun is also present:

(168)   a.   Hebrew

        Gil   cilem   le-Rina$_i$ et   ha-ax/xaver/ima   *(šela$_i$)

        *Gil   photographed   to-Rina   ACC   the-brother/friend/mother *(her)*

        'Gil photographed Rina's brother/friend/mother' (Landau 1999: 14)

        b.   French

        Je   lui$_i$ ai tué   son$_i$/*le frère

        *I   to-her   killed   her/*the brother*

        'I killed her brother' (Landau 1999: 14)

        c.   Spanish

        *Le   llegó   el   padre   a   Daniela

        3SG(DAT)   arrived   the   father   to   Daniela

        'Daniela's father arrived' (María Cristina Cuervo, p.c.)

These generalizations are not applicable to possessive clitics in Pashto. Sentences in which a possessive clitic associates with a relational noun appear above in (143–147), with the relational noun *plaar* 'father' as possessee. The Pashto possessive clitic also may have an external argument as its possessum. As was demonstrated in (27) in chapter 1, the verb *khandel* 'laugh' is

125

unergative: in present tense (169a) it selects a subject in direct case, while in past tense (169b) it selects an oblique (ergative) subject—exactly as if it were a transitive verb—and bears default agreement (masculine 3pl), since the verb never agrees with ergative NPs:

(169)   a.   (ze)   khand-em

         PN1SG(DIR)   laugh-PRES1SG

         'I am laughing'

         b.   maa   khand-él(e)

         PN1SG(OBL)   laugh-PAST(MASC3PL)

         'I laughed'

         c.   *ze   khand-él(e)

         PN1SG(DIR)   laugh-PAST(MASC3PL)

         'I laughed' (Babrakzai 1999: 112; Tegey and Robson 1996: 188)

The same alternation is observed for full NPs that have distinct case forms for direct and oblique roles, as does *wrúna/wrúno* 'brothers (DIR/OBL)' below:

(170)   a.   wrúna   khand-i

         brothers(DIR)   laugh-PRES3SG

         'the brothers are laughing'

         b.   wrúno   khand-el

         brothers(OBL)   laugh-PAST(MASC3PL)

         'the brothers laughed'

                    (Farooq Babrakzai, p.c.; Jan Mohammad, p.c.; Habibullah Tegey, p.c.)

A second-position possessive clitic may be inserted into either of the above sentences. Most strikingly here, in contrast to languages having possessive dative constructions, is that the possessive clitic may associate with the external argument in (171b), which is, moreover, a relational noun 'brother':

(171)   a.   wrúna   <u>mee</u>   khand-i

                brothers(DIR)   *1SG*   *laugh-PRES3SG*

                'my brothers are laughing'

          b.   wrúno   <u>mee</u>   khand-el

                brothers(OBL)   *1SG*   *laugh-PAST(MASC3PL)*

                'my brothers laughed'

                              (Farooq Babrakzai, p.c.; Jan Mohammad, p.c.; Habibullah Tegey, p.c.)

Another unergative verb in Pashto is *ghapel* 'bark'. In the following sentence, the verb is past tense, and so the sole argument appears in its oblique form, *spi*, rather than in its direct form, which would be *spay*. (Unrelated illustrations of the *spi/spay* alternation may be found in (5–8) above and (197–198, 202–203, 263–264) below.)

(172)   begaa   shpa   **dee**   **ahmad**   spi   dder   we   ghap-el

          last   *night*   *POSS*   *Ahmad*   *dog(OBL)*   *very*   *PERF*   *bark-PAST(MASC3PL)*

          'Ahmad's dog barked a lot last night' (Tegey and Robson 1996: 189)

The full NP possessor, 'Ahmad,' may be replaced by a 2P possessive clitic, 3sg *yee*, illustrating once more that—unlike possessive dative constructions in Hebrew—the possessive clitic may associate with an external argument:

(173)   begaa   shpa   <u>yee</u>   spi   dder   we   ghap-el

last   *night*   *3SG*   *dog(OBL)*   *very*   *PERF*   *bark-PAST(MASC3PL)*

'his/her dog barked a lot last night' (Jan Mohammad, p.c.)

Of course, it is equally possible to associate a possessive 2P clitic with the single argument of an unaccusative verb:

(174)   khor   <u>dee</u>   raaghl-a

sister   *2SG*   *came-FEM3SG*

'your sister came' (Babrakzai 1999: 94)

Clearly, the association of Pashto's 2P possessive clitic to a possessee does not distinguish between external and internal arguments; all such associations are equally grammatical, which is distinctly unlike the sharp asymmetries that are displayed by possessive dative constructions in Hebrew.

Landau (1999: 5) teases apart other semantic properties of the possessive dative construction that likewise do not obtain in Pashto's possessive clitic constructions. For example, in genitive constructions with 'picture'-nouns, as in (162a) above, the genitive may be interpreted not only as the possessor of the picture, but also its creator, as well as its theme (i.e., the person depicted by the picture). The corresponding possessive dative construction (162b), however, lacks the theme interpretation for the possessive dative. The same asymmetry obtains in Spanish. No such asymmetry obtains for 'picture' nouns in Pashto, however:

(175)   a.   khoshal   **dee**   **haghee**   'aks   loy   kerr

Khoshal   *POSS*   *PN3SG(FEM)*   *picture*   *big*   *do(PAST PERF MASC3SG)*

'Khoshal enlarged her picture'

128

b. khoshal <u>yee</u> 'aks loy kerr

   Khoshal   3SG   picture   big   do(PAST PERF MASC3SG)

   'Khoshal enlarged her picture' (Jan Mohammad, p.c.)

The same three-way ambiguity as obtains in the Hebrew genitive construction in (162a) obtains in both of the above Pashto sentences, whether the genitive is expressed by a strong pronoun with the preposition *dee* 'of' (175a), or as a 2P clitic *yee* (175b).

Obviously, the possessive dative constructions discussed above do not resemble Pashto's dislocation of second-position possessive clitics. One of the most striking differences is that none of Hebrew, French, and Spanish allows the dative possessor to associate with an external argument, whereas Pashto permits a possessor-possessee relation between a 2P clitic and any argument within its clause.

3.5.2.2. Other possessive alternations

Another variety of possessor raising is found in the Western Muskogean languages, which do permit the displaced possessor to associate with both external and internal arguments.[52] In the following examples from Chickasaw, both the possessor of a subject (176a) and the possessor of an object (177a) may undergo possessor raising, assigning a different grammatical form to the possessor in the corresponding (b) sentences:

(176)   a.   Ihoo   im-ofi'-at   ishto

         woman   AGR(III)-dog-SUBJECT   big

         'The woman's dog is big'

---

[52] The same point has been made by Broadwell (1990: 228), ch. 3 of which discusses Western Muskogean possessor raising in more detail than is relevant here.

b. Ihoo-at ofi'-at im-ishto

woman-SUBJECT *dog-SUBJECT* *AGR(III)-big*

'The woman's dog is big' (Munro 1984: 637)

(177) a. Ofi'-at ihoo im-pask-<u>a</u> apa-tok

dog-SUBJECT *woman* *AGR(III)-bread-NON.SUBJECT* *eat-PERF*

'The dog ate the woman's bread'

b. Ofi'-at ihoo-<u>a</u> pask im-apa-tok

dog-SUBJECT *woman-NON.SUBJECT* *bread* *AGR(III)-eat-PERF*

'The dog ate the woman's bread' (Munro 1984: 637)

Chickasaw is similar to Pashto in that 'possessor raising' may apply either to a subject or an object. However, as Munro (1984) notes, there are differences between subject possessor raising (176) and object possessor raising (177): the altered possessor does not change into the same form in both kinds of sentences, but rather, the possessor of a subject receives a subject affix, while the possessor of an object receives an object (='non-subject') suffix. Moreover, only one non-subject suffix may appear in an object possessor raising sentence, whereas—as shown in (176b)—a subject possessor raising sentence may bear more than one subject suffix. Furthermore, while subject possessor raising derives an otherwise non-occurring sentence type, object possessor raising derives the same form of a sentence that is used for benefactive and other interpretations—thus, (177b) may also be interpreted as 'The dog ate the bread for the woman'.

Several properties of possessor raising in Chickasaw therefore distinguish it from possessive 2P clitics in Pashto. First, in Chickasaw, 'possessor raising' itself appears not to be a unitary phenomena, distinguishing as it does between subjects and objects. Pashto possessive clitics may associate with either subject or object possessees without any observable differences, either in the form of the clitic or the possessee. Second, possessor raising in Chickasaw changes

the possessor into a nominal with a different case form or grammatical function (similar to the possessive dative construction in Hebrew, French, and Spanish). Possessive clitics in Pashto do not appear to have changed grammatical functions. Third, like Hebrew and Romance languages, the possessor raising alternations in Chickasaw involve full NPs. In Pashto, the possessive clitic is, of course, a form that bears only person and number features, and not any lexical information of the sort that is borne by a full NP. Like Hebrew and the other languages, it appears to be tenable to assume for Chickasaw that possessor raising derives structures having arguments (since they are full NPs with argument-like properties discussed in the references above)—the main subject for debate, of course, being exactly what kind of argument has been derived.

Another language with possessor raising is Japanese, as shown by the following alternation.

(178) a. [John-no imooto] -ga utsukusi-i

        John-GEN *sister* *NOM* *beautiful-be*

        'John's sister is beautiful'

    b. John-ga   imooto-ga   utsukusi-i

        John-NOM *sister-NOM* *beautiful-be*

        'John's sister is beautiful' (Ura 1996: 100)

As in Chickasaw and Pashto, possessor raising in Japanese is also possible from the subject of an unergative verb (179a) or a transitive verb (179b):

(179) a. John-ga musuko-ga butai-de odot-ta

        John-NOM *son-NOM* *stage-at* *dance-PAST*

        'John's son danced at the stage'

b.　John-ga　　　musuko-ga　　hito-o　　　　korosi-ta

　　John-*NOM*　　*son-NOM*　　*person-ACC*　　*kill-PAST*

　　'John's son killed a man' (Ura 1996: 108–109)

However, Japanese does not allow possessor raising from an object:

(180)　　a.　John-ga　　　Mary-no　　atama-o　　nagut-ta

　　　　　　John-*NOM*　　*Mary-GEN*　*head-ACC*　*hit-PAST*

　　　　　　'John hit Mary's head'

　　　　　b.　*John-ga　　Mary-o　　atama-o　　nagut-ta

　　　　　　John-*NOM*　　*Mary-ACC*　*head-ACC*　*hit-PAST*

　　　　　　'John hit Mary's head' (Ura 1996: 110)

(181)　　a.　John-ga　　　Mary-no　　kuruma-o　　migai-ta

　　　　　　John-*NOM*　　*Mary-GEN*　*car-ACC*　　*polish-PAST*

　　　　　　'John polished up Mary's car'

　　　　　b.　*John-ga　　Mary-o　　kuruma-o　　migai-ta

　　　　　　John-*NOM*　　*Mary-ACC*　*car-ACC*　　*polish-PAST*

　　　　　　'John polished up Mary's car' (Ura 1996: 110)

In this asymmetry, Japanese possessor raising appears to have the opposite properties of Hebrew. Clearly, possessor raising has strikingly different properties across languages. For our purposes here, however, note that Japanese shares a crucial property with other possessor-raising languages that distinguishes them all from Pashto: the possessive alternations involve full NPs. Pashto does not have such full-NP alternations as exhibited by Japanese in (178–179).

3.5.3. Covert dislocation

The previous section has shown that Pashto's dislocation of the possessive clitic from its possessee is not comparable to superficially similar constructions in other languages. The displaced possessor in Pashto is always a pronominal clitic (not a full NP), and it bears only some features of person and number—exactly the features that reside in agreement projections. Rather than suppose that this clitic is an argument, moved from its base position, it is more perspicuously treated as an agreement marker, and specifically as the head of an agreement projection. The question in (165) posed by Landau (1999) about possessive dative constructions is therefore an inappropriate one to ask concerning Pashto's displaced possessive clitics, as they do not bear any resemblance to arguments.

Landau (1999: 9) adopts a case-driven movement analysis of possessive dative constructions, with the following properties:

(182)   a.   The possessor is generated in a caseless Spec position within the possessee

b.   It is generated with dative case features

c.   It then raises [to Spec/VP] to check its case features with V

It is obvious that this analysis is inappropriate for Pashto 2P possessive clitics. The first problem with applying the analysis in (182) to Pashto is that possessive 2P clitics do not bear dative case features. Whereas the 'raised' possessor in Hebrew, French, and Spanish, appears in dative case, the 2P possessive clitic in Pashto does not bear any resemblance whatsoever to a dative NP. There is a separate set of dative clitics that appear near the verb (see section 3.4.5 above), but the 2P clitics themselves fill functions that (other than genitive) are best described as ergative and accusative (i.e., oblique). Second, as was suggested in chapter 2, Spec/VP is typically occupied by an adjective or noun that serves as the first constituent of compound verbs—and the bulk of Pashto's verbs are compound, rather than simple—rendering Spec/VP generally unavailable as a

target for NP-movement. Third, having the dative possessor move to Spec/VP in Hebrew explains why a possessive dative may not move out of an external argument: it would entail downward movement from the external argument position (Spec/*v*P) to the more deeply embedded Spec/VP position, an option excluded under Minimalist assumptions (Landau 1999: 11–12). However, as was shown above, Pashto freely allows possessor dislocation from external arguments.

The movement aspect of the analysis in (182) may be maintained, though, without requiring the problematic overt downward movement, if the actual argument is the empty category *pro*. The clitic is a clausal agreement morpheme generated as the head of an Agr projection, where it appears to be in 'second position'. As an empty category, *pro* does not move for case reasons, contra (182), but rather moves in order to erase features of person and number— intuitively, to be 'identified' as far as the grammar of Pashto permits, i.e., as far as the person/number combinations that are spelled out as clitics.

Recall from the list of pronominal 2P clitics in (82) that they may function as ergative, accusative, or genitive.[53] It is impossible to know, when citing a single clitic out of context, whether it is ergative, accusative, or genitive, since it may have any of those functions. Pashto clitics are thus distinctly unlike weak pronouns in other languages having pronominal clitics, whether they are verbal clitics, as in Spanish, or second-position clitics, as in Serbo-Croatian. As may be seen from the following inventories of pronominal clitics in those languages, their forms alone often (in third-person, always) suffice to distinguish their grammatical function as dative or accusative (and, for Serbo-Croatian, also genitive):

---

[53] It might be possible to collapse ergative and genitive into a single category, given the subject-like properties of genitives (Chomsky 1970, Alexiadou and Wilder 1998), but it is difficult at present to see how the accusative function of these clitics might also be subsumed under this category.

(183)   Spanish weak pronouns (Perlmutter 1971: 20)

		Dative	Accusative
1sg		me	me
2sg		te	te
3sg	masc	le	lo
	fem	le	la
1pl		nos	nos
3pl	masc	les	los
	fem	les	las
3sg & pl	reflexive	se	se

(184)   Serbo-Croatian second-position pronominals (Schütze 1994: 376)

		Dative	Genitive	Accusative
1sg		mi	me	me
2sg		ti	te	te
3sg	masc/neut	mu	ga	ga
	fem	joj	je	je/ju[54]
reflexive		si	se	se
1pl		nam	nas	nas
2pl		vam	vas	vas
3pl		im	ih	ih

Another respect in which Pashto provides a notable typological contrast to these languages is in its ordering of clitics. Recall from the template in (86), repeated below as (185), that Pashto's pronominal 2P clitics are ordered according to grammatical number, with first- and second-person preceding third-person:[55]

---

[54] Although the genitive and accusative forms are almost entirely identical, even they may be distinguished by their 3sg feminine forms, since the 3sg feminine accusative clitic *je* surfaces as *ju* when the otherwise homophonous 3sg auxiliary clitic *je* is also present in the cluster (Schütze 1994: 420).

[55] Ch. 4 shows that this ordering is derived directly from the syntax, rather than from a morphological template.

(185)

1	2	3	4
kho	ba	am	am
			mo
'indeed'	'will'	1PL, 2PL	1PL, 2PL

5	6	7	8
mee	dee	yee	no
1SG	2SG;	3SG, 3PL	'then'
	'should'		

The order schematized in (185) differs from the relative order of pronominal 2P clitics in Spanish and Serbo-Croatian, both of which are ordered according to their case (dative preceding accusative). The following template is generally assumed for Serbo-Croatian (Schütze 1994: 375; Tomic 1996: 816):

(186)   Interrog > Aux (except 3sg) > Dative > Acc/Gen > Reflexive > Aux (3sg)

Pashto's clitics are not ordered by their case—according to the analysis advanced here, because the arguments they identify bear only a single case: oblique. But even observationally, Pashto's clitics are not ordered according to their functions, as will be explicated below. As the first step in dispensing with the template in (185), it should be noted that it is misleading, insofar as it suggests that all eight slots could be filled in a single sentence. In fact, the number of pronominal clitics in a sentence is determined by the number of arguments selected by the verb. Tegey (1977) established the above template, pairwise, from such sentences as the following; only the first interpretations, numbered (i), are from Tegey; the other interpretations will be discussed below:

(187)   a.   topak   <u>mee</u>   <u>dee</u>   raaworr-e

                   gun(MASC)   1SG   2SG   brought-MASC3SG

                   (i)   'I brought your gun' (Tegey 1977: 185)

                   (ii)   'You brought my gun'

136

b.   topak   <u>mee</u>   <u>yee</u>   raaworr-e

gun(MASC)   *1SG*   *3SG*   *brought-MASC3SG*

(i)   'I brought his gun' (Tegey 1977: 185)

(ii)   'He brought my gun'

c.   topak   <u>dee</u>   <u>yee</u>   raaworr-e

gun(MASC)   *2SG*   *3SG*   *brought-MASC3SG*

(i)   'You brought his gun' (Tegey 1977: 185)

(ii)   'He brought your gun'

d.   topak   <u>am</u>   <u>mee</u>   raaworr-e

gun(MASC)   *1PL*   *1SG*   *brought-MASC3SG*

(i)   'I brought our gun' (Tegey 1977: 185)

(ii)   'We brought my gun'

e.   topak   <u>am</u>   <u>dee</u>   raaworr-e

gun(MASC)   *2PL*   *2SG*   *brought-MASC3SG*

(i)   'You (pl) brought your (sg) gun' (Tegey 1977: 185)

(ii)   'You (sg) brought your (pl) gun'

f.   topak   <u>am</u>   <u>yee</u>   raaworr-e

gun(MASC)   *1PL*   *3SG*   *brought-MASC3SG*

(i)   'We brought his gun' (Tegey 1977: 185)

(ii)   'He brought our gun'

g. topak   <u>am</u>   <u>mo</u>   raaworr-e

    gun(MASC)   *1/2PL*   *1/2PL*   *brought-MASC3SG*

    (i)   'We brought your gun' (Tegey 1977: 185)

    (ii)   'You brought our gun'

    (iii)   'We brought our gun'

    (iv)   'You (PL) brought your (PL) gun'

It is tempting to suppose from this paradigm that the first clitic is interpreted as the ergative, and the second as the genitive (Babrakzai 1999: 94); however, the above sentences are ambiguous in their interpretation, as indicated by the alternative glosses listed in (ii) for each example.[56] This point was not made clear by Tegey (1977: 185), who intended merely to establish the relative order of 2P clitics with respect to their form, not to suggest that the first such clitic receives an ergative interpretation.[57] Rather, as is usual for weak and strong pronouns in any language, their reference is variable, but fixed by principles of binding and discourse, so that ambiguity does not normally arise.

Indeed, it is difficult to demonstrate the ordering schematized in (185) with a single sentence, as a sentence containing more than two pronominal clitics is difficult to parse, lacking even a dominant reading (Jan Mohammad, p.c.). The point is amply illustrated by the following pair of sentences. While (188a), with two pronominal clitics, is already two-ways ambiguous, merely adding a third pronominal clitic—intended as a possessor within an adjunct, as in (188b)—increases the ambiguity considerably, not even considering the singular/plural ambiguity of third-person *yee*:

---

56 Thanks to Jan Mohammad for confirming the ambiguity of these sentences.

57 See Tegey (1977: 182, ex. 3) for a similar example of four-way ambiguity.

(188) a. plaar   kho   ba   mee   dee   no   ne-byaay-i

         father   *indeed*   FUT   *1SG*   *2SG*   *then*   NEG-*bring*-PRES3SG

        (i)   my father is not going to take you along, then?

        (ii)   your father is not going to take me along, then?

    b. plaar   kho   ba   mee   dee   yee   no

         father   *indeed*   FUT   *1SG*   *2SG*   *3SG/PL*   *then*

        pe   motar   ki   ne-byaay-i

        in   *car*   *in*   NEG-*bring*-PRES3SG

        (i)   my father is not going to take you along in his car, then?

        (ii)   my father is not going to take him along in your car, then?

        (iii)   your father is not going to take me along in his car, then?

        (iv)   your father is not going to take him along in my car, then?

        (v)   his father is not going to take me along in your car, then?

        (vi)   his father is not going to take you along in my car, then?

                                        (Jan Mohammad, p.c.)

As indicated by the glosses in (188b), each clitic may be construed as the object, as the possessor for 'father', and as the possessor for 'car.' Such sentences are strongly preferred to have at least one strong pronoun, as the pronoun would have to be clearly associated with one or the other of these grammatical functions, bringing the ambiguity of (188b) back to the more manageable ambiguity in (188a). The sentences in (188) are therefore reminiscent of such notoriously difficult sentences as the following, which involve center-embedding:

(189) a. That the fact that he left was unfortunate is obvious. (Miller and Chomsky 1963: 471)

    b. The rat the cat the dog chased ate died. (Newmeyer 1986: 172)

The grammar permits such sentences, but they are frequently so difficult to parse as to be unusable.

Having demonstrated the ambiguity of Pashto's 2P pronominal clitics, the analysis that immediately suggests itself is the same as one that seems reasonable for explaining the cases of non-3rd-person weak pronouns in Spanish (183) and non-3rd-person genitive and accusative 2P pronouns in Serbo-Croatian (184); namely, to assume that homophonous forms are nevertheless distinctly case-marked, parallel to other pronouns in the paradigm whose forms show such case-marking overtly (e.g., in English, compare the distinctly marked 3rd-person nominative *he/she* vs. accusative *him/her*, to the 2nd-person *you*, which is both nominative and accusative). This sort of analysis could be extended to Pashto, and it would entail dividing the pronominal 2P clitics of (82) into three sets—ergative, accusative, and genitive—all of which happen to be identical.

Such an analysis, though, would miss an important generalization, in that the homophony of these clitics would be regarded as accidental. If the pronominal 2P clitics of Pashto are instead regarded as a single, coherent set of agreement morphemes, a generalization can be captured—that their forms are identical—and their varying functions may be predicted on the basis of independently required principles of verb agreement and thematic structure. Thus, unlike (strong) pronouns in English—which are assumed to be arguments, often distinguishing case by their form alone—and unlike strong pronouns in Pashto (which occur in fixed argument positions, and distinguish two cases), Pashto's 2P clitic pronouns are ambiguously interpreted exactly because they are associated with a single case, oblique, which is assigned to three functions (ergative, accusative, genitive); in this regard, then, the clitics behave distinctly unlike arguments, which show a two-way case-distinction. On the other hand, a property shared by all of Pashto, Spanish, and Serbo-Croatian is that their clitic pronouns do not bear a nominative role. In Pashto, the lack of a nominative/absolutive clitic is expected under the assumption that clitics represent agreement, since nominative/absolutive agreement is morphologically represented elsewhere in the language: as a suffix on the verb. Conversely, there is no ergative/accusative/genitive

agreement suffix in Pashto because this category of agreement is independently represented by the 2P clitics.

Let us now consider in more detail how possessive *pro* is licensed, examining the ambiguous interpretations that may arise at LF. Consider again the derivation of sentences whose clitics are ambiguously interpreted, examples of which were given in (146) and (187). The sentence in (187a) is repeated as (190) below.

(190)   topak   <u>mee</u>   <u>dee</u>   raaworr-e

           gun   *1SG*   *2SG*   *brought-3SG*

           (i)   'I brought your gun' (Tegey 1977: 185)

           (ii)   'You brought my gun'

This ambiguity arises for two reasons: first, pronominal clitics bear a single case, oblique, which is associated with three different functions (ergative, accusative, genitive); and second, the clitics have a fixed order with respect to each other, as was schematized in the template in (185). Making the first move to dispense with the template in (185)—which will be pursued further still in the next chapter—we are led to embedded Clitic Phrases (which, recall, are simply AgrPs). Before Spell Out, both interpretations of (190) have the following structure, with the object *topak* 'gun' in its base position:

(191)

```
                    CliticP
              /              \
       mee 1SG            CliticP
                      /            \
                dee 2SG             vP
                              /           \
                          pro₁             VP
                                      /          \
                                    DP            V
                                /       \          |
                            pro₂         NP      raaworre
                                        /  \     'brought'
                                      topak 'gun'
```
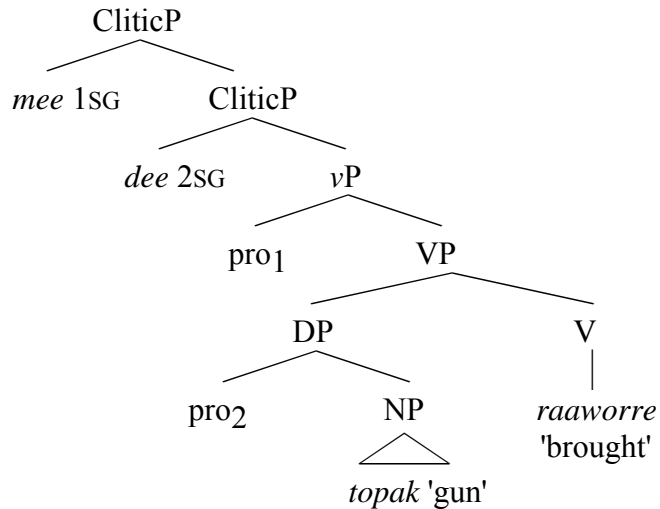
There being no full NP subject, the direct object may satisfy the EPP by moving to Spec/TP:

(192)

```
                         TP
                  /             \
               DPᵢ              CliticP
             /    \          /           \
         pro₂      NP    mee 1SG          CliticP
                  /  \               /            \
                topak 'gun'       dee 2SG           vP
                                            /           \
                                         pro₁            VP
                                                    /        \
                                                  tᵢ          V
                                                              |
                                                           raaworre
                                                           'brought'
```
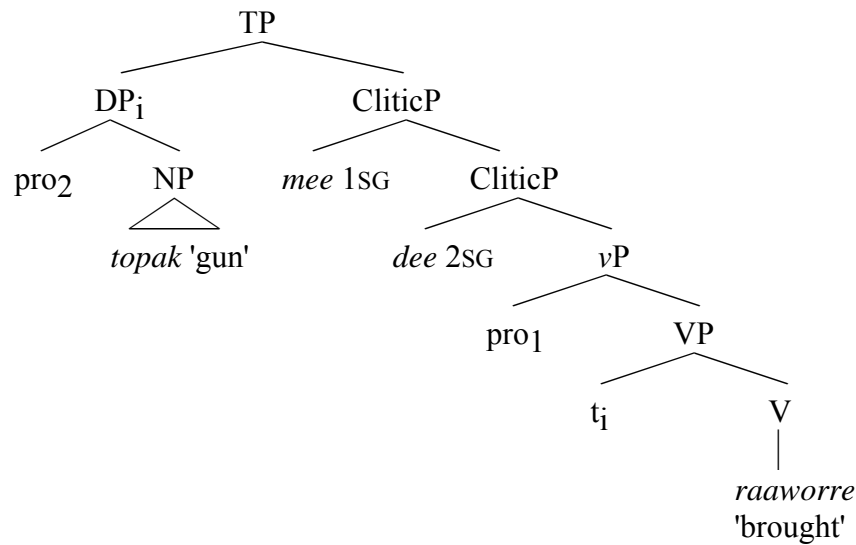
There are two options for erasing the person and number features in CliticP: either pro_1 or pro_2 moves to specifier position of the higher CliticP, headed by 1sg *mee,* and the other *pro* moves to specifier position of the lower CliticP, headed by 2sg *dee.* The choice between these options is free, deriving the two interpretations indicated in (190).

Some technical questions arise concerning this derivation, specifically concerning the movement of pro_2 (the possessive pronominal inside the direct object). While its interpretation

142

as a possessor is derived from its base position in Spec/DP, its movement is unusual. Either pro_2 moves to Spec/CliticP at the stage of derivation illustrated in (191)—with the stranded part of the direct object, *topak* 'gun', moving later to Spec/TP, resulting in more-or-less successive steps of movement—or else the full direct object (pro_2 gun) moves to Spec/TP first, as shown in (192), with pro_2 then moving to Spec/CliticP. The latter movement would have to be downward, while both options otherwise are not strictly successively cyclic, as they require extraction of pro_2. As noted above with regard to (103), though, cyclicity does not constrain LF operations, and so if *pro* moves only at LF (i.e., only after the representation in (192) has been derived by overt syntax), general principles of well-formedness are obeyed. The only strong consideration for the movement of *pro* to Spec/CliticP at LF is that there be the right number of such projections to accommodate the empty pronominals that need to be identified; and conversely, that there be enough empty pronominals to erase the person and number features of any clitics that may be present. As mentioned earlier, this issue is tightly and independently constrained by the argument structure of the verb.

Another striking consequence of this analysis of possessor *pro* is that it requires that the agreement relation between *pro* and the person/number features are not within NP/DP (where *pro* is projected/merged), but rather are within the clause, TP. Although the analysis of clitics suggested here leads us into this conclusion, it is important to bear in mind that while it results in an unexpected relation of agreement, the surprise is not so much with the theory as it is with the language itself. Even a descriptive account of Pashto clitics needs to state that while the possessive clitics are interpreted at the site of a nominal element, their placement is determined within the clause. It is this property of 'dislocation' from their site of interpretation that makes clitic pronouns different from their strong pronoun counterparts. And given that this unusual relation actually obtains in the language, it is best to express it by covert movement at LF (i.e., by treating clitics as agreement), rather than by movement of the clitic itself. As was shown in (159–161), Pashto does not tolerate overt movement of possessive NPs.

3.5.4. Locality of interpretation

We have seen that when a transitive sentence has a single full NP and two clitics, the clitics are ambiguous as to where they are interpreted: one must be the object, while the other must be the possessor of the overt NP, whether the overt NP is the subject or the object. These interpretations are determined by the verb's argument structure, while the ambiguities are determined by the fixed ordering of the single-cased clitics with respect to each other. Despite appearances until now, however, clitics are not always ambiguously interpreted. It is possible to construct sentences in which the interpretation of the clitic is determined by its syntactic position with respect to surrounding nominals; in such cases, the possessive clitic must be interpreted with the highest NP in the sentence.

In a transitive sentence having two full NPs and a single clitic, one overt NP is the subject, while the other is the object. The verb's argument structure having been saturated by the full NPs, the remaining, single clitic must be interpreted as genitive, although it might be expected to be ambiguous as to whether it associates with the subject or the object. In fact, this is not the case. The possessive clitic only associates with the higher overt argument, regardless of whether it is the subject (in SOV order) or the object (in OSV order):

(193)　　a.　kheza　　　mee　　kitaab　　　lwel-i

　　　　　　　wife(DIR)　　*1SG*　　*book(DIR)*　　*read-PRES3SG*

　　　　　　　(i)　　'my wife is reading a book'

　　　　　　　(ii)　*'the wife is reading my book'

　　　　　b.　kitaab　　　mee　　kheza　　　lwel-i

　　　　　　　book(DIR)　　*1SG*　　*wife(DIR)*　　*read-PRES3SG*

　　　　　　　(i)　　'the wife is reading my book'

　　　　　　　(ii)　*'my wife is reading a book'

<div align="right">(Farooq Babrakzai, p.c.; Jan Mohammad, p.c.)</div>

(194)   a.   khezee   <u>mee</u>   kitaab   lwest-e

wife(OBL)   1SG   book(DIR)   read(PAST)-MASC3SG

(i)   'my wife was reading a book'

(ii)   *'the wife was reading my book'

b.   kitaab   <u>mee</u>   khezee   lwest-e

book(DIR)   1SG   wife(OBL)   read(PAST)-MASC3SG

(i)   'the wife was reading my book'

(ii)   *'my wife was reading a book'

(Farooq Babrakzai, p.c.; Jan Mohammad, p.c.)

That the clitic appears to be attached to its possessum has nothing to do with its interpretation. Adding an adverb, *maaspekhin* 'afternoon', to the beginning of these sentences coaxes the possessive clitic away from its possessee, into second position, and yet the same interpretation obtains, with the possessive clitic taking the nearest NP as its possessee:

(195)   a.   maaspekhin   <u>mee</u>   kheza   kitaab   lwel-i

afternoon   1SG   wife(DIR)   book(DIR)   read-PRES3SG

(i)   'my wife is reading a book in the afternoon'

(ii)   *'the wife is reading my book in the afternoon'

b.   maaspekhin   <u>mee</u>   kitaab   kheza   lwel-i

afternoon   1SG   book(DIR)   wife(DIR)   read-PRES3SG

(i)   'the wife is reading my book in the afternoon'

(ii)   *'my wife is reading a book in the afternoon'

(Farooq Babrakzai, p.c.)

145

(196)   a.   maaspekhin   <u>mee</u>   khezee   kitaab   lwest-e

afternoon   *1SG*   *wife(OBL)*   *book(DIR)*   *read(PAST)-MASC3SG*

(i)   'my wife was reading a book in the afternoon'

(ii)   *'the wife was reading my book in the afternoon'

b.   maaspekhin   <u>mee</u>   kitaab   khezee   lwest-e

afternoon   *1SG*   *book(DIR)*   *wife(OBL)*   *read(PAST)-MASC3SG*

(i)   'the wife was reading my book in the afternoon'

(ii)   *'my wife was reading a book in the afternoon'

(Farooq Babrakzai, p.c.)

The same asymmetry obtains in sentences containing NPs that denote potentially reciprocal participants, such as 'cat' and 'dog' in (197–198) below—though note that in order to remove potential confusion as to which NP is the agent in present tense, a heavy pause is required after the fronted object (indicated by a comma) in OSV order in (198b), which also requires that the clitic seem to appear in 'third position' (although, of course, it is in 'second position' of its own clause, the object being topicalized). In the past tense sentences of (197), there is ergative case-marking and agreement, and so both SOV and OSV orders are available:

(197)   a.   spi   <u>mee</u>   pisho   khog   krr-a

dog(OBL MASC)   *1SG*   *cat(DIR FEM)*   *hurt*   *do(PAST PERF)-FEM3SG*

(i)   'my dog hurt the cat'

(ii)   ?'the dog hurt my cat'

b.   pisho   <u>mee</u>   spi   khog   krr-a

cat(DIR FEM)   *1SG*   *dog(OBL MASC)*   *hurt*   *do(PAST PERF)-FEM3SG*

(i)   'the dog hurt my cat'

(ii)   ?'my dog hurt the cat'

(Jan Mohammad, p.c.)

(198)   a.   spay   <u>mee</u>   pisho   khog-aw-i

dog(DIR)   *1SG*   *cat(DIR)*   *hurt-TRANS(PRES IMPF)-3SG*

(i)   'my dog is hurting the cat'

(ii)   ?'the dog is hurting my cat'

b.   spay,   pisho   <u>mee</u>   khog-aw-i

dog(DIR)   *cat(DIR)*   *1SG*   *hurt-TRANS(PRES IMPF)-3SG*

(i)   'my cat is hurting the dog'

(ii)   'the dog, he is hurting my cat'

(iii) *'the cat is hurting my dog'

(Jan Mohammad, p.c.)

The striking point of these data is that scrambling, an overt syntactic operation, feeds the interpretation of clitics at LF: the possessive clitic takes the highest NP as its possessum, irrespective of tense or word order. At Spell Out, the structure of the OSV sentence of (197b) above would be as follows, with the object (*pisho* 'cat') and its possessor *pro* having moved to Spec/TP to erase the EPP feature (or else adjoined or merged in a topic/focus position):

(199)

```
                        TP
                  ┌─────┴─────┐
                DPᵢ          CliticP
              ┌──┴──┐      ┌────┴────┐
            pro   pisho  Clitic'
                  'cat'  ┌────┴────┐
                       Clitic      vP
                         │      ┌───┴───┐
                      mee 1SG  NP      VP
                              spi 'dog' ┌──△──┐
                                      tᵢ khog krra 'hurt do'
```
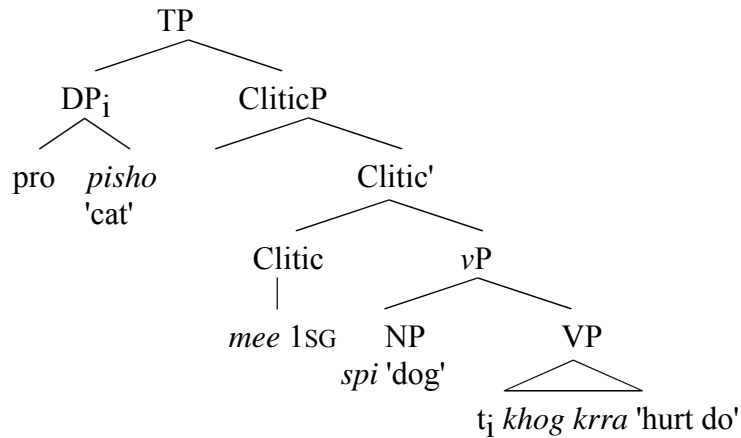
The head of CliticP has person and number features (1sg) that need to be checked at LF. If the derivation proceeded from bottom-up, we would expect the nearest NP, *spi* 'dog' to move to Spec/CliticP to erase the features:

(200)

```
                      * TP
                 ┌─────┴─────┐
               DPᵢ          CliticP
             ┌──┴──┐     ┌────┴─────┐
           pro  pisho  NPₖ       Clitic'
                'cat'  spi 'dog' ┌───┴───┐
                              Clitic     vP
                                │     ┌───┴───┐
                             mee 1SG tₖ      VP
                                          ┌──△──┐
                                        tᵢ khog krra 'hurt do'
```
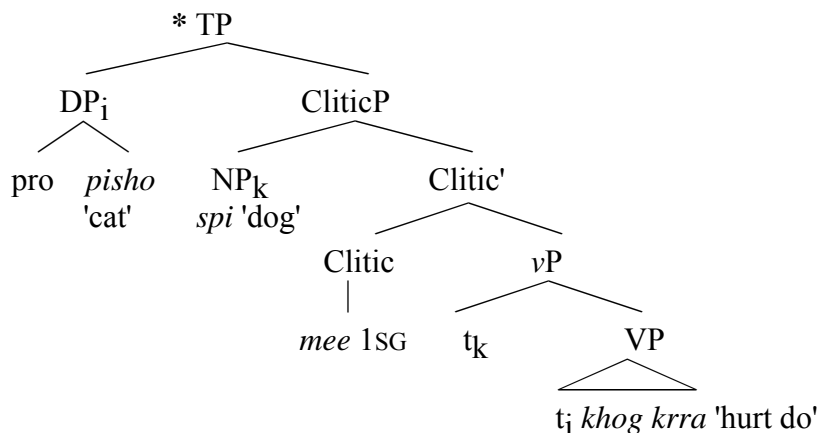
This derivation crashes for two reasons. First, there is a mismatch in person features between the moved NP (3sg) and the clitic head (1sg). Second, possessive *pro* (of *pisho* 'cat') fails to be interpreted. Another option in deriving the LF representation from (199) is to move the highest DP containing the possessor *pro* and *pisho* 'cat' to Spec/CliticP, but this too would result in a number-feature mismatch (3sg vs. 1sg)—aside from the consideration that it would derive a clitic-doubled element, which 2P clitics do not do. The only remaining option, then, is to move the possessor alone:

(201)

```
                    TP
             _____|_____
           DP_i            CliticP
          __|__          ____|____
         t_k  pisho    pro_k     Clitic'
              'cat'           ____|____
                           Clitic        vP
                             |         ___|___
                          mee 1SG    NP       VP
                                   spi 'dog'   /\
                                            t_i khog krra 'hurt do'
```
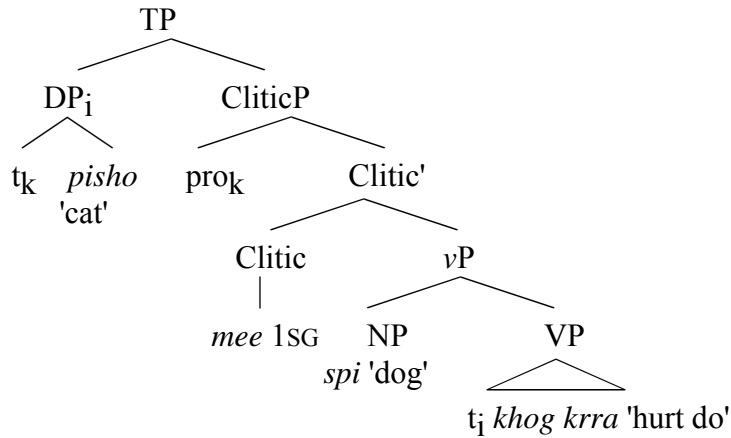
This derivation is licit because cyclicity does not obtain at LF, and because movement of the empty category *pro* does not violate the Left Branch Condition (151), which in Pashto only constrains overt movement; cf. (159–161) above.

The sentence in (198b) is of further interest, because it shows that the genitive clitic is clause-bound in its interpretation, and may not associate with a dislocated NP. The interpretation in (i) is the OSV order, while the interpretation in (ii) is subject-dislocation, in which case the 1sg clitic *mee* is obliged to function as possessor of 'cat' in its clause. The same point may be illustrated by having a heavy pause after the initial NP in the other sentences above. As in (198b), these sentences require the clitic to be in apparent 'third position', and they would be ungrammatical without the heavy pause (indicated by a comma) after the fronted NP:

(202)   a.   spi,   pisho   <u>mee</u>   khog   krr-a

               dog(OBL MASC)   cat(DIR FEM)   1SG   hurt   do(PAST PERF)-FEM3SG

               (i)   'the dog hurt my cat'

               (ii)   *'my dog hurt the cat'

b. pisho,   spi   <u>mee</u>   khog   krr-a

cat(DIR FEM)   dog(OBL MASC)   1SG   *hurt   do(PAST PERF)-FEM3SG*

(i)   'my dog hurt the cat'

(ii)   *'the dog hurt my cat'

(203)   spay,   pisho   <u>mee</u>   khog-aw-i

dog(DIR)   cat(DIR)   1SG   hurt-TRANS(PRES IMPF)-3SG

(i)   'the dog is hurting my cat'

(ii)   *'my dog is hurting the cat'

(Jan Mohammad, p.c.)

Although when a clitic is obliged to be possessive it must be interpreted with the highest NP in its clause, the grammar retains the possibility of other options, if forced into them. For example, if the highest overt argument happens not to be able to take a possessor (as when the argument is a pronoun or a proper name), then the clitic takes the lower overt NP as its possessee:

(204)   a.   te   <u>mee</u>   plaar   sara   khaberi   wu   krr-a

PN2SG   1SG   father   with   news   PERF   do(PAST PERF)-IMP2SG

'you talk to my father!'

b.   Tor   <u>mee</u>   plaar   meelma   kerr

Tor   1SG   father   guest   DO(PAST PERF MASC3SG)

'Tor invited my father/my father invited Tor' (Jan Mohammad, p.c.)

There is therefore a locality constraint on clitic interpretation, but it may be overridden by the stronger requirement for the clitic to be interpreted at LF.

150

Interpretation of the possessive clitic is yet more constrained than simply being clause-bound. Specifically, if one or more NPs is contained within a larger NP, the clitic may only be interpreted with the head of the larger NP, as illustrated below:

(205) a. dee   khowúnki   yaadúna   yee   khe   di

POSS   teacher   memories(MASC DIR)   3SG   good(MASC)   be(PRES3PL)

(i)   'his memories of the teacher are good'

(ii)   *'memories of his teacher are good'

b. dee   khowúnki   dee   wror   dee   kor

POSS   teacher   POSS   brother   POSS   house

yaadúna   mee   khe   di

memories(MASC DIR)   1SG   good(MASC)   be(PRES3PL)

(i)   'my memories of the house of the brother of the teacher are good'

(ii)   *'memories of the house of the brother of my teacher are good'

(Jan Mohammad, p.c.)

The subject of (205a) has the following structure, in which *pro* in Spec/DP represents the possessor that is associated with the 3sg clitic *yee* in the matrix clause:

(206)

The possessor *pro* cannot instead be located in Spec/DP of the lower nominal, *khowúnki* 'teacher',
because this position would be too deeply embedded to allow *pro* to move to Spec/CliticP without
violating locality.

3.6. Relative clauses

Further evidence that clitics are agreement, not arguments, comes from an asymmetry in relative
clause (RC) formation. Relativization of nominative/absolutive arguments results in a gap in the
RC, whereas relativization of ergative, accusative, and genitive arguments results in a 2P clitic
appearing within the RC. While such clitics co-occur with the relative *wh*-word, this appearance
of *wh*-movement and clitic doubling (the appearance of an argument and a coreferential
pronominal clitic within the same clause) will be shown to be illusory.

3.6.1. Gapping asymmetries

In English, no matter what position is relativized (subject, object, possessor), there is a gap in the
RC where the argument would normally appear in a non-RC, and which is the position from
which the relative *wh*-pronoun (assumed to be the argument) has been moved; resumptive
pronouns are ungrammatical in those positions, as demonstrated by the (b) sentences below:

(207)   Subject

           a.   Fred met the guy [who$_i$ t$_i$ ate all the pizza]

           b.   *Fred met the guy [who$_i$ he$_i$ ate all the pizza]

(208)   Object

           a.   Fred met the guy [who$_i$ Mary likes t$_i$]

           b.   *Fred met the guy [who$_i$ Mary likes him$_i$]

(209)   Possessor

    a.   Fred met the guy [whose father$_i$ Mary saw t$_i$]

    b.   *Fred met the guy [whose father$_i$ Mary saw {him/his$_i$}]

The pattern in Pashto is different: while English has a gap in all (subjacent) relativized positions, Pashto has a gap in some RCs, but a 2P clitic pronoun in others. As noted by Babrakzai (1999: 108), relativized nominatives (210) and absolutives (211)—which constitute a single class, as may be seen from their identical (direct) case-marking—have a gap inside the RC:[58]

(210)   Nominative

maa	agha	dzhiney	welid-a
I(ERG)	*DET*	*girl(ABS)*	*saw-FEM3SG*

[tshee	∅	pe	kaabal	kee	oseeg-i]
COMP		*at*	*Kabul*	*in*	*live-3SG*

'I saw the girl who lives in Kabul' (Babrakzai 1999: 109)

(211)   Absolutive

maa	agha	spay$_i$	welid-e
I(ERG)	*DET*	*dog(ABS)*	*saw-MASC3SG*

[tshee	taa	mor	ta	∅$_i$	werkerray	w-e]
COMP	*you(ERG)*	*mother*	*to*		*give*	*was-MASC3SG*

'I saw the dog which you had given to mother' (Babrakzai 1999: 109)

---

[58] In the following examples, ∅ marks the site of extraction from the (bracketed) relative clause, and it is coindexed with its intended referent. Only the internal syntax of RCs (i.e., the grammatical role of the extracted element within the RC itself) is considered in the examples presented here, since it reveals the relevant asymmetry regarding agreement. For discussion of the external syntax of RCs (i.e., the function of the head of the RC with respect to the matrix clause), see Tegey (1977: ch. 4; 1979: 392 ff.).

It is also an option to include a *wh*-word, which then precedes the complementizer, as shown in (212–213)—RCs being the only environment in which Pashto appears to have *wh*-movement. Notice that a pronominal clitic marking the location of the gap inside the RC is ungrammatical:

(212)   Nominative

maa   agha   dzhiney   welid-a   [(sok)   tshee

I(ERG)   det   girl(ABS)   saw-FEM3SG   who(NOM)   COMP

pe   kaabal   kee   (*yee)   wuseeg-i]

at   Kabul   in   3SG   live-3SG

'I saw the girl who lives in Kabul' (Laghman)

(213)   Absolutive

maa   agha   spay   welid-e   [(kem)   tshee

I(ERG)   det   dog(ABS)   saw-MASC3SG   which   COMP

taa   (*yee)   mor   ta   werkerray   w-e]

you(ERG)   3SG   mother   to   give   was-MASC3SG

'I saw the dog which you had given to mother' (Laghman)

In contrast, relativized accusatives (214), ergatives (215–216), and genitives (217) have a 2P pronominal clitic inside the RC, coreferring with the head of the RC. The accusative RC in (214) contains the 3rd-person clitic pronoun, *yee*; see also (114) above for discussion of clitic placement in this sentence.

(214)   Accusative

agha   kisa   [tshee   dzhan   <u>yee</u>   lwel-i]

DET   *story(FEM SG)*   *COMP*   *John*   *3SG*   *read-PRES3SG*

ddeera   ugda   da

very(FEM SG)   *long(FEM SG)*   *be(FEM3SG)*

'The story that John is reading is very long' (Babrakzai 1999: 108)

As noted above, relativized ergatives also have an ergative clitic inside the RC. Additionally, however, the ergative RC head must appear in absolutive form:[59]

(215)   Ergative

agha   kheza   [tshee   kitab   <u>yee</u>   waa-khist-e]

det(ABS)   *woman(ABS)*   *COMP*   *book(ABS)*   *3SG*   *PERF-take-MASC3SG*

paysee   <u>yee</u>   maa   ta   raakrr-ee

money(ABS)   *3SG*   *I(OBL)*   *to*   *gave-FEM3PL*

'That woman who took the book gave me the money' (Babrakzai 1999: 109)

(216)   Ergative

*aghee   khezee   [tshee   kitab   <u>yee</u>   waa-khist-e]

det(ERG)   *woman(ERG)*   *COMP*   *book(ABS)*   *3SG*   *PERF-take-MASC3SG*

paysee   <u>yee</u>   maa   ta   raakrr-ee

money(ABS)   *she*   *I(OBL)*   *to*   *gave-FEM3PL*

'That woman who took the book gave me the money' (Babrakzai 1999: 109)

---

[59] Note that when the head of the RC is a transitive subject in past tense (i.e., ergative), a coreferential pronominal clitic appears in the matrix clause, as well as in the RC itself. This is only one property of the RC's external syntax that led Tegey (1977: ch. 4) to suggest that relative clauses and their heads are obligatorily left-dislocated—surely the correct analysis, given the evidence about the interpretation of clitics to be presented in the next section.

Relativized genitives also have a clitic at the site of the RC gap:

(217)   Genitive

agha   kheza   [tshee   wror   <u>yee</u>   taarikh   lwel-i]

det(ABS)   woman(ABS)   COMP   brother   3SG   history   study-PRES3SG

zmaa   gaawenddey   da

PN1SG(POSS)   neighbor(FEM)   be(PRES IMPF FEM)

'That woman whose brother is studying history is my neighbor'

(Babrakzai 1999: 110)

While Babrakzai (1999: 108) remarks that the clitic is obligatory (see also (114a) above for the same observation in another variety of Pashto), in other varieties of Pashto the clitic is merely optional:

(218)   Accusative

agha   kisa   [(kem)   tshee   dzhan   (<u>yee</u>)   lwel-i]

DET   story(FEM SG)   which   COMP   John   3SG   read-PRES3SG

ddeera   ugda   da

very(FEM SG)   long(FEM SG)   be(FEM3SG)

'The story that John is reading is very long' (Laghman)

(219)   Ergative

agha   kheza[60]   [(tshaa)   tshee   kitab   (yee)

det(ABS)   *woman(ABS)*   *who(ERG)*   *COMP*   *book(ABS)*   *3SG*

waa-khist-e]   paysee   yee   maa   ta   raakrr-ee

PERF-take-MASC3SG   *money(ABS)*   *3SG*   *I(OBL)*   *to*   *gave-FEM3PL*

'That woman who took the book gave me the money' (Laghman)

Nevertheless, even this variety of Pashto retains the crucial contrast between the possibility of having an RC-internal clitic in accusatives (218) and ergatives (219), and its impossibility in nominatives (212) and absolutives (213).

Two properties of relative clauses are therefore puzzling: (i) they appear to have *wh*-movement, and (ii) they induce apparent clitic doubling in some environments. Neither of these properties is displayed in *wh*-questions. Like Hindi and Japanese, Pashto is a *wh-in-situ* language (i.e., questions do not involve overt *wh*-movement); not only does the *wh*-word of a question remain *in situ*, but there is no appearance of clitic doubling:

(220)   a.   asad   kar   kawi

                  Asad   *work*   *do*

                  'Asad is working'

          b.   asad   tse   kawi

                  Asad   *what*   *do*

                  'What is Asad doing?'

---

[60] In spoken Laghman Pashto, it is possible for *aghee khezee* 'that woman (ERG)' to head the RC.

(221)　　a.　laylaa　　tshalaw　　pakhawi

　　　　　　　Layla　　*rice*　　　*cook*

　　　　　　　'Layla is cooking the rice'

　　　　　b.　tsok　　　　tshalaw　　pakhawi

　　　　　　　who(DIR)　　*rice*　　　*cook*

　　　　　　　'Who is cooking the rice?'

(222)　　a.　ágha　dee　　laylaa　kitaab　day

　　　　　　　that　*POSS*　*Layla*　*book*　*be*

　　　　　　　'That's Layla's book'

　　　　　b.　ágha　dee　　tshaa　　　kitaab　day

　　　　　　　that　*POSS*　*who(OBL)*　*book*　*be*

　　　　　　　'Whose book is that?' (Tegey and Robson 1996: 168–169)

It will be suggested that the asymmetry between relative clauses and questions in this regard reflects an unusual property not of the questions, but of the RCs: when a *wh*-word appears in an RC, it is left-dislocated—specifically, merged directly in its surface position, to the left of the complementizer.[61] Because the *wh*-word does not originate in the RC itself, the RC does not contain a gap (*wh*-trace), but instead an empty, resumptive pronominal (*pro*).

　　　Evidence that the gap in the RC is *pro* is that clitics only appear in relative clauses in which they would also appear in a corresponding main clause, i.e., to identify ergative, accusative, and genitive arguments. This asymmetry in the presence (or absence) of a clitic inside

---

[61] As noted above, similar facts led Tegey (1977: ch. 4; 1979: 392 ff.) to this conclusion also for the external syntax of the RCs, which he argued are derived by left-dislocation of both the head of the RC and the RC itself.

the RC is strong evidence that these clitics constitute agreement morphemes, not arguments, since the split is between nominative/absolutive (gap) versus ergative/accusative/genitive (clitic). While it is common for languages to display a restriction on relativization of ergatives, a restriction on ergative/accusative/genitive is more unusual, if not actually unattested. With respect to the head of the RC, Pashto does indeed display the familiar restriction on relativization of ergatives: as was shown in (215–216) above, an ergative RC must appear in absolutive form in order for it to be relativized. With respect to pronominal clitics inside the RC, however, there is no reason to suppose that their presence reflects any special restriction on the internal structure of the RC itself, since the clitics that appear in such RCs are restricted to the single (oblique) case that identifies these same three functions in a main clause lacking a full NP or strong pronoun argument: ergative, accusative, and genitive. In the relativized nominatives (210) and absolutives (211), the verb inside the RC shows the usual agreement with the extracted element, and so a clitic would never be expected in those environments, anyway—indeed, the clitic was shown to be ungrammatical in the corresponding sentences of (212–213). In the relativized accusative (214), however, the present tense verb inside the RC agrees with the subject (as it would in a main clause), and so object agreement in this environment is also manifested as it would be in a main clause, i.e., as a clitic.

A close inspection of the behavior of clitics suggests that, despite appearances, Pashto has neither clitic doubling, nor *wh*-movement in relative clauses. Recall the relativized accusative in (218) above, repeated below:

(223)   agha   kisa   [(kem)   tshee   dzhan   (yee)   lwel-i]

          DET   *story(FEM SG)*   *which*   *COMP*   *John*   *3SG*   *read-PRES3SG*

          ddeera   ugda   da

          very(FEM SG)   *long(FEM SG)*   *be(FEM3SG)*

          'The story that John is reading is very long' (Laghman)

Consider the possibility that Pashto is like English in having *wh*-movement in relative clauses. Assuming that the fronted *wh*-word *kem* 'which' has been moved from the argument position inside the RC, the relative clause portion of (223) would have the following structure (focusing only on the form of the relative clause):

(224)

```
                          NP
            ┌──────────────┴──────────────┐
           NP                             CP
        ╱──────╲                   ┌───────┴───────┐
     agha kisa                    NP              C'
     DET story                    │           ┌────┴────┐
                                 kemᵢ         C         TP
                                which         │      ┌───┴───┐
                                            tshee    NP    CliticP
                                            COMP     │    ┌────┴────┐
                                                   dzhan  t2ᵢ     Clitic'
                                                   John        ┌────┴────┐
                                                             Clitic      VP
                                                               │      ┌───┴───┐
                                                              yee     NP      V
                                                              3SG     │    lwel-i
                                                                     t1ᵢ   read-PRES3SG
```
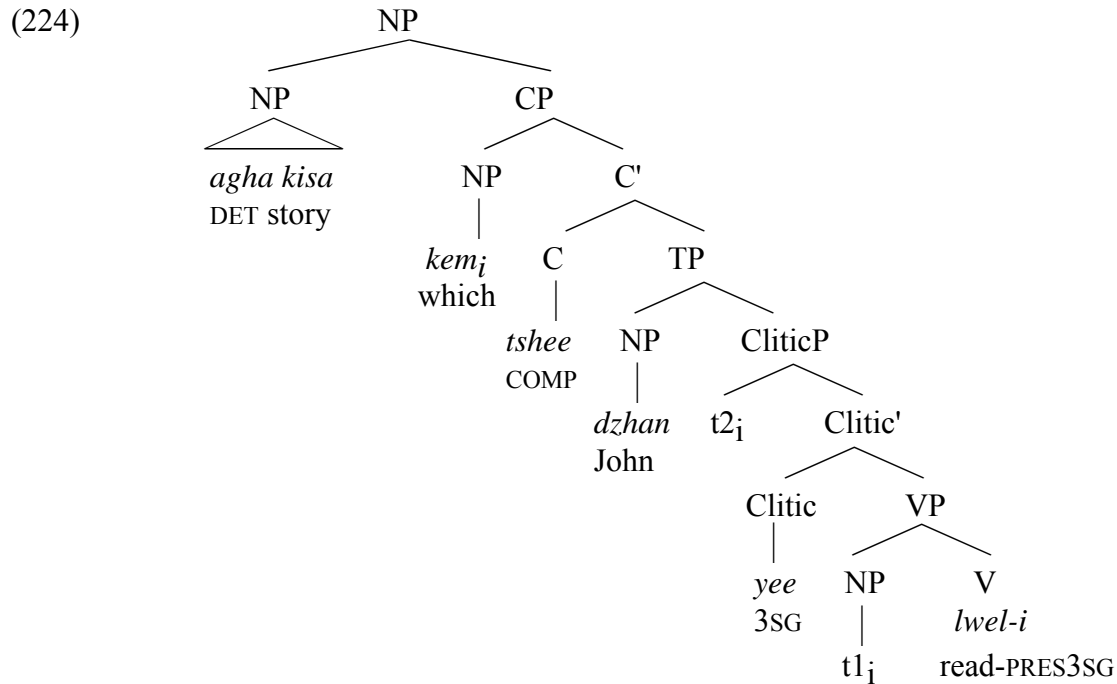
In the above structure, the subject, 'John', is shown in its surface position, Spec/TP, though its base position is actually Spec/*v*P (not shown here). The direct object, *kem* 'which', would be the argument of the verb, originates at the site of t_1. On its way to Spec/CP (via *wh*-movement), it substitutes in Spec/CliticP (the site of t_2), where it checks the 3sg feature.

Relative clauses of this kind, containing clitics to identify the *wh*-trace, offer a striking contrast to structures like (103) above, in which *pro* is licensed in Spec/CliticP. In particular, the licensing of *pro* in (103) may be seen as serving the purpose of identifying the empty category, in the same way that verbal agreement suffixes license *pro* drop. This sort of explanation collapses, though, in trying to explain why a clitic appears in certain relative clauses, since the contents of the empty category (trace) ought to be clear enough, due to the presence of both the displaced *wh*-

160

word and the head of the relative clause—'the story' in (224)—which is coindexed with the *wh*-word. Positing a *wh*-trace in relative clauses would therefore result in two undesirable asymmetries. The first asymmetry is that clitics would be the heads of agreement projections that identify disparate categories: *pro* and *wh*-trace. The only property shared by these empty categories is their phonological form, which is null. The second asymmetry is that clitic doubling occurs only with *wh*-words in relative clauses, in which the *wh*-word is displaced from the position in which it is interpreted, whereas ordinary *wh*-questions have neither *wh*-movement nor clitic doubling.

The best way to remove these asymmetries is to regard the apparently moved *wh*-word in relative clauses as left-dislocated: the *wh*-word is merged directly in its surface position, and so the gap inside the RC itself is *pro*, which is identified either by a clitic head or by verbal agreement, as in any ordinary main clause seen until now. Consider again the relative clause in (219). The ergative *wh*-word *tshaa* 'who' appears to have been moved, since it appears to the left of the complementizer *tshee*. The suggestion now is that the *wh*-word actually appears in a left-dislocated position, as illustrated below in (225). The object, *kitab* 'book', is in Spec/TP, pursuant to the discussion in section 3.4.2.2 above. The clause from which the ergative *wh*-word *tshaa* 'who' appears to have been extracted therefore contains not *wh*-trace, but rather *pro*, as projected by the verb in Spec/*v*P and moved to Spec/AgrP in order to be identified by the 3sg features associated with the 2P clitic *yee*. This *pro* is coindexed with the *wh*-word, which is presumably in an external focus position, rather than in Spec/CP via *wh*-movement. Although the *wh*-word has ergative case, this cannot be taken to indicate that it has moved from the RC; see (257) below (and the note there) for an example of a left-dislocated ergative NP.

(225)

```
                    FocusP
                   /      \
            tshaa_i         CP
            who(ERG)       /  \
                          C    TP
                        tshee  /  \
                       COMP  NP    AgrP
                            kitab  /   \
                         book(ABS) pro_i  Agr'
                                         /   \
                                       Agr    vP
                                     yee 3SG  /\
                                        ... waakhiste 'took'
```
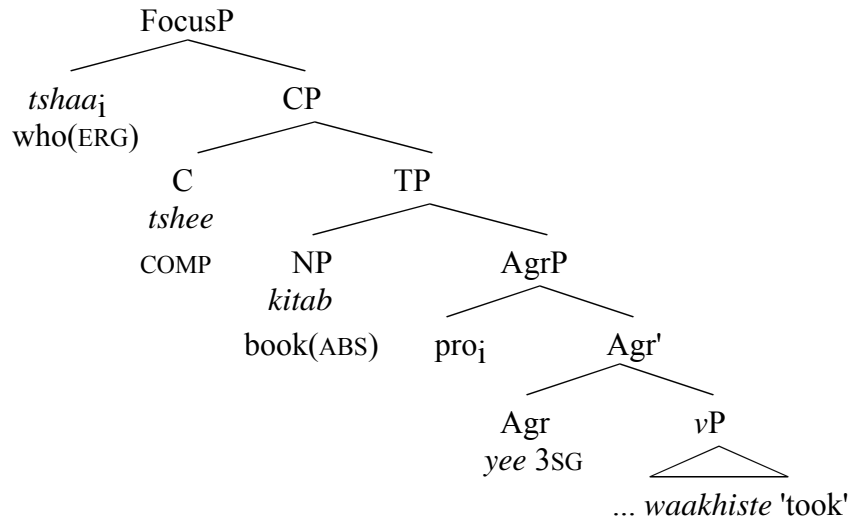
This analysis permits a simple generalization about the appearance of pronominal clitics: they only serve to identify *pro*. This proposal removes the otherwise unusual asymmetry in relative clauses—in which only ergative-, accusative-, and genitive-centered relative clauses contain a clitic rather than a gap—because the appearance of a clitic inside an RC has nothing to do with the fact that it is a relative clause; clitics now appear in the same places in both main and relative clauses, identifying *pro*. The structure in (225) lacks a variable, however, which is required for the relative clause to be interpreted. The next section will show that this variable is created not by movement (as in English and other European languages), but by null-operator binding of *pro* (as in similar constructions in Hebrew and Irish). In other words, *pro* may function as a resumptive pronoun.

3.6.2. Resumption and dislocation

There is independent evidence that the empty category (EC) inside the RC is pronominal, rather than a trace, which can be uncovered as we try to determine the nature of the empty category inside the relative clause. Sells (1984) distinguishes resumptive pronouns from intrusive pronouns: aside from differences in their interpretation to be discussed below, the former have the distribution of gaps and do not improve island violations, whereas the latter may repair island violations, perhaps because they are the spell-out of the illicit trace (Shlonsky 1992; Pesetsky

1997, 1998). The following English sentences exemplify intrusive pronouns. (226a) violates the Left Branch Constraint (151), while (226b) relativizes from within a relative clause, violating the Complex NP Constraint. For speakers who permit intrusive pronouns, they may be used to improve such island violations:

(226)   a.   All the students who the papers which *(they) submitted were lousy I'm not going to allow to register next term.

            b.   The only kind of car which I can never get *(its) carburetor adjusted right is them Stanley Steamers. (Ross 1967: 260–261)

Intrusive pronouns must be distinguished from resumptive pronouns, which appear in positions in which a gap would be expected. Hebrew and Irish are well known for their resumptive pronouns:

(227)   a.   ze   ha'iš   [še   **oto**   ra'iti   etmol]

                this-is   the man   COMP   him   I-saw   yesterday

                'This is the man that I saw yesterday' (Sells 1984: 6)

            b.   an   rud   [aN   gcoinníonn tú   ceilte   orthu   **é**]

                the   thing   COMP   keep(PRES)   you   concealed   on-them   it

                'the thing that you keep concealed from them'

                                     (McCloskey and Hale 1983: 497)

Unlike the intrusive pronouns in (226), the resumptive pronouns *oto* 'him' in Hebrew and *é* 'it' in Irish appear here in subjacent relative clauses. The literal translations of (227a,b) in English are, of course, ungrammatical, since English does not have resumptive pronouns—only intrusive pronouns, which are restricted to islands:

(228)   a.   This is the man that I saw (*him) yesterday

          b.   the thing that you keep (*it) concealed from them

Because the pronominal clitic in Pashto appears at the site of the gap within subjacent relative clauses, Pashto appears more like Hebrew and Irish than English.

An examination of islands in Pashto suggests that the EC identified by a clitic is indeed a resumptive pronoun (i.e., *pro*), rather than an intrusive pronoun (i.e., the spell-out of a *wh*-trace). Some sentences that would be weak island violations in English are fully grammatical in Pashto:

(229)   maa   hagha   tsok   tshi   ne-poh-ig-em

          PN1SG(OBL)   that   who(DIR)   COMP   NEG-wise-INTR-1SG

          Tor   meelma   kérr-ey   wu

          Tor   guest   do(PAST PERF)-PART   be(PAST3SG)

          (aw)   ka   na   we-lid-e

          and   or   not   PERF-saw-MASC3SG

          'I saw the person who I don't know whether Tor had invited or not'

                                                                    (Jan Mohammad, p.c.; Yusufzai)

However, a strong island violation is created by relativizing from within another relative clause, and the presence of the 3sg (ergative) clitic *yee* at the most deeply embedded extraction site does not improve its grammaticality (as does an intrusive pronoun in the English gloss):

(230)　　*hágha　　sarray　　[tshee　　kema　　dzhorra

　　　　　　DET　　　*man(DIR)*　　*COMP*　　*which*　　*suit*

　　　　　[tshee　　(<u>yee</u>)　　aghust-ay　　wa]　　　　　　genda　　wa]

　　　　　 COMP　　*3SG*　　*wear-PART*　　*be(PAST.IMPF3SG)*　　*dirty*　　*be*

　　　　　deer　　bad　　khkar-id-o

　　　　　much　　*bad*　　*look-INTR-PAST3SG*

　　　　　'the man who the suit that (he) was wearing was dirty looked very bad'

　　　　　　　　　　　　　　　　　　　　　　　　(Yusufzai; Jan Mohammad, p.c.)

Nor may a strong pronoun at the extraction site serve as an intrusive pronoun. The following sentence is as ungrammatical as the previous one:

(231)　　*hágha　　sarray　　[tshee　　kema　　dzhorra

　　　　　　DET　　　*man(DIR)*　　*COMP*　　*which*　　*suit*

　　　　　[tshee　　**agha**　　aghust-ay　　wa]　　　　　　genda　　wa]

　　　　　 COMP　　*PN3SG*　　*wear-PART*　　*be(PAST.IMPF3SG)*　　*dirty*　　*be*

　　　　　deer　　bad　　khkar-id-o

　　　　　much　　*bad*　　*look-INTR-PAST3SG*

　　　　　'the man who the suit that (he) was wearing was dirty looked very bad'

　　　　　　　　　　　　　　　　　　　　　　　　　　(Jan Mohammad, p.c.)

Note also that the ungrammaticality of (230) is not due to the fact that the 3sg clitic *yee* immediately follows the complementizer *tshee*; see (112–113) above for discussion of grammatical sentences in which the clitic follows the complementizer. The fact that the clitic pronoun in (230) may not serve as an intrusive pronoun for the purpose of improving an island violation, then, suggests that in ordinary (specifically, ergative-, accusative-, and genitive-centered) relative clauses, the appearance of such a clitic also does not serve an intrusive

function; in other words, the clitic pronoun may not be considered the spell-out of a *wh*-trace. In the context of the analysis here, in which clitics are agreement, this conclusion is not surprising, and constitutes further evidence that clitics do not behave as arguments.

The islands above contain an ergative gap. The same subjacency effect obtains when the gap is nominative—a position in which clitics never appear:

(232)   *hágha   sarray$_i$ [tshi   hágha   deerishi

        DET   *man(DIR)*   *COMP*   *DET*   *suit(FEM DIR)*

        [tshi   Ø$_i$ <u>yee</u>   ághund-i]   khirena   da]

        COMP   *3SG*   *wear-PRES3SG*   *dirty(FEM DIR SG)*   *be(PRES IMPF FEM3SG)*

        nen   deer   bad   khkar-ig-i

        today   *much*   *bad*   *look-INTR-PRES3SG*

        'the man who the suit that (he) wears is dirty looks very bad today'

                                                          (Jan Mohammad, p.c.)

A strong pronoun at the extraction site improves the sentence only slightly:

(233)   ?hágha   sarray   [tshi   hágha   deerishi

        DET   *man(DIR)*   *COMP*   *DET*   *suit(FEM DIR)*

        [tshi   **day/agha**   <u>yee</u>   ághund-i]   khirena   da]

        COMP   *PN3SG(VIS/INVIS)*   *3SG*   *wear-PRES3SG*   *dirty*   *be(PRES IMPF3SG)*

        nen   deer   bad   khkar-ig-i

        today   *much*   *bad*   *look-INTR-PRES3SG*

        'the man who the suit that he wears is dirty looks very bad today'

                                                           (Jan Mohammad, p.c.)

The fact that both ergative and nominative gaps behave identically with respect to island violations suggests that the empty category identified by clitics and verbal agreement suffixes is indeed one and the same, i.e., *pro*.

Because island violations have not hitherto been constructed for Pashto, it is useful to show that the constituent parts of the sentences above are grammatical. The constituents of the ergative-centered RC in (230) are as follows:

(234)　　　hágha　　sarray　　nen　　deer　　bad　　khkar-id-o

　　　　　　DET　　*man(DIR)*　*today*　*much*　*bad*　*look-INTR-PAST3SG*

　　　　　　'the man looked very bad today'

(235)　　　hágha　dzhorra　tshee　kem　sarri　　aghust-ay　wa　genda　wa[62]

　　　　　　DET　*suit*　　*COMP*　*DET*　*man(OBL)*　*wear-part*　*was*　*dirty*　*was*

　　　　　　'the suit that the man was wearing was dirty'

　　　　　　　　　　　　　　　　　　　　　　　　　　(Yusufzai; Jan Mohammad, p.c.)

The constituents of the nominative-centered RC in (232) are as follows:

(236)　　　hágha　　sarray　　nen　　deer　　bad　　khkar-ig-i

　　　　　　DET　　*man(DIR)*　*today*　*much*　*bad*　*look-INTR-PRES3SG*

　　　　　　'the man looks very bad today'

---

[62] The phrase *kem sarri* is best translated as 'the man', although *kem* more familiarly means 'some; which'. The usage here is typical of Peshawar and, possibly, Eastern Afghanistan (Jan Mohammad, p.c.).

(237)　　　hágha　deerishi　　　tshi　　sarray　　　<u>yee</u>　ághund-i

　　　　　　DET　　*suit(FEM DIR)*　　*COMP*　　*man(DIR)*　　*3SG*　　*wear-PRES3SG*

　　　khirena　　　　　　da

　　　dirty(FEM DIR SG)　　*be(PRES IMPF FEM3SG)*

　　　'the suit that the man wears is dirty'

<div align="right">(Jan Mohammad, p.c.)</div>

That clitics do not improve island violations, however, does not entirely suffice to show that they are resumptive rather than intrusive. Igbo, for example, is similar to Pashto in that both gaps and resumptive pronouns obey island constraints (Sells 1984: 213, citing Goldsmith 1981). Rather, the hallmark way in which resumptive pronouns and intrusive pronouns differ is in their interpretations: a resumptive pronoun is interpreted as a bound variable, while an intrusive pronoun may never be thus interpreted. Consider the ordinary pronoun *he* in (238), which may have either of the interpretations in (239):

(238)　　　Only John likes the girl **he** is dancing with

(239)　　　a.　　Bound variable

　　　　　　　　Only John is an x such that x likes the girl that **x** is dancing with

　　　　　　b.　　Referential

　　　　　　　　Only John is an x such that x likes the girl that **John** is dancing with

<div align="right">(Sells 1984: 7–8)</div>

The truth conditions of these interpretations differ, as may be seen in a context like the following. Suppose there are only two men and their partners. If Fred doesn't like his own partner, but does

like John's partner, the bound variable interpretation (a) will be true, and the referential interpretation (b) will be false.

Sells (1984) claims that resumptive pronouns only have the bound variable interpretation, since they have the same distribution as gaps (which are bound by an operator). The gap in subjacent relative clauses in English, for example, receives a bound variable interpretation:

(240)   a.   the man that Bill saw __

          b.   the man x such that Bill saw x

However, it is not the case that English pronouns can never have a bound variable interpretation; as shown above in (238), they may indeed do so. Rather, pronouns in English can have the bound variable interpretation only when they are bound by a quantificational phrase in an A-position, such as 'only John' above, or 'every man' below:

(241)   a.   Every man thinks that Mary likes him

          b.   Every man x thinks Mary likes x

Sells calls this 'anaphor binding,' to distinguish it from operator binding. While all languages have the former relation, only languages like Hebrew and Irish (and, as we shall see, Pashto) also allow pronouns to be operator-bound. Notice, for example, that the English intrusive pronoun is permitted in the non-quantificational relative clause (242a), but forbidden in the quantificational relative clause in (242b):

(242)   a.   I'd like to meet the linguist [that Mary couldn't remember if she had seen **(him)** before]

b. I'd like to meet every linguist [that Mary couldn't remember if she had seen (***him**) before] (Sells 1984: 11–12)

English differs from Hebrew in this regard. As shown below, the resumptive pronoun in Hebrew may be quantificationally bound from an operator position:

(243) kol gever še dina xoševet [še **hu** ohevet rina]

every man COMP Dina thinks COMP he loves Rina

'every man that Dina thinks loves Rina' (Sells 1984: 16)

Taking these binding asymmetries as the defining characteristics of resumptive pronouns, Sells (1984: 27) proposes the following definitions:

(244) a. A pronoun that is interpreted as a bound variable whose antecedent is in an A-position is **anaphorically bound**

b. A pronoun that is interpreted as a bound variable whose antecedent is an operator is a **resumptive pronoun**

c. A pronoun whose antecedent is in an A'-position but which is not interpreted as a bound variable is an **intrusive pronoun**

With respect to the above diagnostics, the pronominal clitics in Pashto's relative clauses behave as resumptive pronouns. Exactly as in Hebrew (243)—and unlike English (242)—the clitic inside the relative clause may be bound by an empty operator coreferential with a quantificational head:

(245)   Ergative

ar   sarray   [tshi   te   <u>yee</u>   milma

every(MASC)   man(MASC DIR)   COMP   PN2SG(DIR)   3SG   guest

kérr-ey   wee]   laarr

do(PAST PERF)-PART   be(PAST IMPF 2SG)   go(PAST PERF.MASC3SG)

'every man who had invited you left' (cf. (115a) above)

(246)   Accusative

ára   kisa   [tshi   dzhan   <u>yee</u>   lwel-i]

every(FEM)   story(FEM DIR)   COMP   John   3SG   read-PRES3SG

ddeera   ugda   da

very(FEM SG)   long(FEM SG)   be(FEM SG)

'every story that John is reading is very long' (cf. (114a), (214), (218) above)

(Jan Mohammad, p.c.)

The clitic that appears obligatorily inside such relative clauses in Pashto is therefore taken to be a true resumptive pronoun, rather than an intrusive pronoun of the kind that exists in English. To be more precise: because the clitic is an agreement head that identifies *pro* in its specifier, it is *pro* that is the resumptive pronoun (not the clitic itself); the clitic merely serves to identify (agree with) *pro*, as has been argued to occur in main clauses. See also (257–258) below for examples of quantifier left-dislocation that also support this point.

The data above therefore constitute additional evidence for the structure of relative clauses proposed in (225): when a *wh*-word appears at the front of a relative clause, it is not an argument that has been moved from the relative clause to Spec/CP (as happens in English), but rather the *wh*-word has been merged directly in the initial position (presumably a focus projection). Taking into account that the clitic inside the RC identifies resumptive *pro* (which is bound by a null operator), the structure of the relative clause in (219) above is not (225), but more fully as in

(247). The direct arguments (absolute *kitab* 'book' and ergative *pro*) originate inside VP, but they are shown here in their Spell Out positions, above VP:

(247)

```
                        FocusP
                       /      \
                 tshaa         CP
               who (ERG)      /   \
                           Op_i    C'
                                  /   \
                                 C     TP
                                 |    /   \
                              tshee  NP    AgrP
                              COMP  kitab  /    \
                                    book pro_i  Agr'
                                               /    \
                                             Agr     AspP
                                              |      /    \
                                             yee   waa-khiste
                                             3sg   PERF-take
```
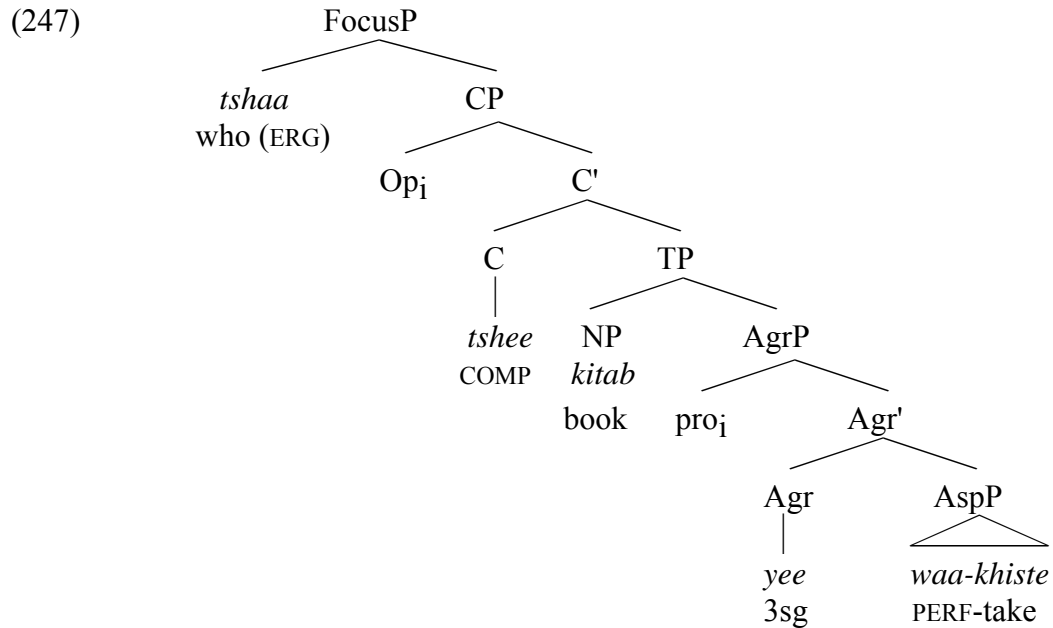
Although there is no *wh*-trace inside relative clauses, this does not pose a problem for interpretation, since *pro* is resumptive (A'-bound by the null operator in Spec/CP), and thereby receives the bound variable interpretation required of a relative clause, exactly as in Hebrew and Irish.

Additional evidence in support of the idea that the fronted *wh*-word of an RC is left-dislocated, and that there is no A'-bound trace inside the relative clause, is that clitics appear to 'double' left-dislocated elements, but may not serve as a reconstruction site—presumably because the left-dislocated element is merged directly in its surface position, rather than being moved there. Recall that because verbs in present tense mark their arguments alike with direct case, a fixed SOV order obtains in sentences having agentively reciprocal participants (e.g., 'cat' and 'dog'). In such cases, OSV order is possible only by following the fronted object with a heavy pause. The following sentences are repeated from (7) and (8) in chapter 1:

172

(248)   a.   Topicalization

          spay,   pisho   khog-aw-i

          dog(DIR)   *cat(DIR)*   *hurt-TRANS(PRES IMPF)-3SG*

          'the cat is hurting the dog'

          b.   Left-dislocation

          spay,   pisho   <u>yee</u>   khog-aw-i

          dog(DIR)   *cat(DIR)*   *3SG*   *hurt-TRANS(PRES IMPF)-3SG*

          'the dog, the cat is hurting him' (Jan Mohammad, p.c.)

That the left-dislocated argument in (248b) is merged in its surface position may be demonstrated by fronting an anaphor. As was shown in (117b)—repeated below as (249a)—the resulting sentence is grammatical only when the anaphor is topicalized; a left-dislocated anaphor is ungrammatical:

(249)   a.   Topicalization

          khpel   zaan   spay   khog-aw-i

          own   *self*   *dog(DIR)*   *hurt-TRANS(PRES IMPF)-3SG*

          'the dog is hurting himself'

          b.   Left-dislocation

          *khpel   zaan,   spay   <u>yee</u>   khog-aw-i

          own   *self*   *dog(dir)*   *3sg*   *hurt-trans(pres impf)-3sg*

          'the dog is hurting himself' (Jan Mohammad, p.c.)

Topicalization of the anaphor in (249a) is fully grammatical even without a pause after the fronted object—despite the identical case-marking on the two NPs in this sentence—since the

173

object may be straightforwardly construed as the anaphor (which the topicalized, non-anaphoric object in (248a) may not be, and which therefore requires a heavy pause following it). The patterns above are identical to those in English, for speakers who permit these constructions (though with full pronouns, of course, rather than a clitic, as in Pashto):[63]

(250)   a.   Fred$_i$, Bill likes t$_i$

        b.   Fred, Bill likes him

(251)   a.   Himself$_i$, Bill likes t$_i$

        b.   *Himself, Bill likes him

This paradigm is handled straightforwardly if topicalization involves movement, as indicated by the traces in the (a) sentences. In (251a), the anaphor has moved to an A'-position, from which it may reconstruct at LF in order to be locally bound by the subject. Because the anaphor in the left-dislocated sentence in (251b) has been merged in its surface position, however, it does not have the option of reconstructing to the position of the object pronoun, and hence is ungrammatical due to Condition A (the anaphor is unbound).[64]

      The same explanation extends to the relevant Pashto sentences above. Specifically, the pronominal clitic only appears to be doubling the dislocated element because they happen to corefer; but the pronominal clitic, argued to head an agreement phrase, invariably licenses *pro* in its specifier. The ungrammaticality of (249b) is due to the fact that the anaphor has not moved from the clause containing the verb, hence cannot lower into it at LF, because the desired

---

[63] See van Riemsdijk (1997) for similar examples in Dutch.

[64] Hindi (which is closely related to Pashto) also distinguishes topicalization from left-dislocation, and the separate derivations of these constructions are surely the same in both languages: topicalization is derived by movement of the NP (adjoining it to TP), while left-dislocation involves merging the dislocated NP in Spec/TopicP, which is external to CP (Dwivedi 1994: 28–29).

reconstruction site is already occupied by *pro*, as shown by the appearance of the clitic there: clitics do not identify traces, only *pro*. The implication of this for relative clauses is that the clitic that appears in certain RCs invariably licenses resumptive *pro*, rather than *wh*-trace. Under this analysis, the fact that clitics appear only in ergative-, accusative-, and genitive-centered relative clauses is explained: in main clauses it is likewise only those functions that take the form of oblique clitics, identifying *pro*. Verbal agreement also identifies *pro*, but bears the complementary direct case (which covers nominative and absolutive roles).

It should be noted that it is only Pashto's full anaphor *khpel zaan* that needs to reconstruct in order to be bound. Unlike (249b), the possessive reflexive may be left-dislocated:

(252)　　　khpéla　　　　　　　ghwaa,　　　　Tor　　yee　　kharts-aw-i

　　　　　　self(POSS FEM DIR SG)　*cow(FEM SG)*　*Tor*　*3SG*　*spent-TRANS(PRES IMPF)-3SG*

　　　　　　'his cow, Tor is selling her' (Jan Mohammad, p.c.)

Whatever the reason for this asymmetry, a similar asymmetry obtains in English. Alongside the ungrammatical (251b) is the following:

(253)　　　His$_i$ (own) father, Bill$_i$ likes him

Presumably, coreference obtains in these sentences not under compulsion from binding principles—although the principles do not exclude the coreference, since the possessive pronoun in both cases does not bind the R-expression.

Further evidence that the Pashto clitic identifies resumptive *pro* is that a quantificational NP may be left-dislocated from the clitic. Recall from (244b) above that pronouns having a bound variable reading are resumptive when their binder is in an A'-position. As noted by Demirdache (1997: 197–198) concerning English, a quantificational phrase may be topicalized but not left-dislocated:

(254)   Topicalization

          every man$_i$, Fred had invited t$_i$

(255)   Left-dislocation

          a.   *every man$_i$, he$_i$ had invited you

          b.   *every story$_i$, John is reading it$_i$

This asymmetry illustrates, again, that topicalization is A'-movement, which creates the gap that is necessary for the bound variable interpretation. Left-dislocation, on the other hand, does not involve movement, and the intrusive pronoun may not be operator-bound. (English does not have resumptive pronouns.) As the sentences in (255) do not have a variable, they violate a general ban on vacuous quantification. In contrast, languages like Hebrew and Egyptian have resumptive pronouns, which means they may be operator-bound:

(256)   a.   kol   gever$_i$,   Rina   xoševet   'alav$_i$

                every   man   Rina   thinks   about-him
                'every man, Rina thinks about him'

          b.   kull   'ustaaz$_i$ mašhuur,   laylaa   'ablit-uh$_i$
                every   professor   famous   Laila   met-him
                'every famous professor, Laila met him' (Demirdache 1997: 198)

With respect to this diagnostic, Pashto patterns like Hebrew and Egyptian. As shown below, it is fully grammatical to left-dislocate quantificational phrases from a clitic:

(257)   Ergative

ar   sarri,   te   <u>yee</u>   milma   kérr-ey   wee[65]

every(MASC)   man(MASC OBL)   PN2SG(DIR)   3SG   guest   did-PART   was(2SG)

'every man, he had invited you'

(258)   Accusative

ára   kisa,   dzhan   <u>yee</u>   lwel-i

every(FEM)   story(FEM DIR)   John   3SG   read-PRES3SG

'every story, John is reading it'

(Jan Mohammad, p.c.)

The appearance of a pronominal clitic in certain relative clauses, then, is taken as evidence that the co-occurring *wh*-word is in a dislocated position. This fact is obvious, too, from the word order, since the *wh*-word precedes the complementizer, well outside the clause from which it would otherwise appear to have been extracted. The appearance of *wh*-movement in some relative clauses is therefore illusory. A pronominal clitic in RCs appears in the usual (matrix) functions exactly because such clauses contain a null pronominal (*pro*), rather than the trace of *wh*-movement. As a *wh-in-situ* language, Pashto does not countenance *wh*-traces in overt syntax. The variable that is required by a relative clause is created through null-operator binding of resumptive *pro*.[66]

---

[65] The direct case form of the left-dislocated NP, *ar sarray*, is ungrammatical here, showing that (inherent) oblique case may be borne by left-dislocated ergative NPs.

[66] An alternative analysis would permit *wh*-movement in RCs: clitics could identify case-marked, A'-bound traces, which would include *wh*-trace and the trace of *pro*—the latter having moved as a topic *à la* Huang (1984). The interpretation of clitics in non-RCs as topics initially suggests the correctness of such an analysis. However, as shown in the text, clitics do not serve as a reconstruction site, making it unlikely that they identify an A'-trace. A further problem with such an analysis would be that it would introduce an unexplained asymmetry between relative clauses and *wh*-questions: the former would require movement, while the latter would exclude it. The

The next chapter turns to the ordering of clitics within the cluster, showing that—just as their appearance in second position of the clause is due to their being merged directly in that position—the order of clitics with respect to each other is also explained by having the syntax merge them directly into their surface position. Moreover, only the pronominal 2P clitics have been discussed so far, and the following chapter will integrate the non-pronominal 2P clitics into this analysis.

---

analysis in the text covers the facts much more simply: clitics invariably identify *pro* (which may be resumptive, i.e., operator bound), and *wh*-traces are uniformly excluded.

4. Ordering within the cluster

4.1. Deriving the template

The preceding chapter showed that the appearance of pronominal clitics in second-position of their clause may be derived without syntactic movement of the clitics themselves, if the clitics are regarded as agreement morphemes (rather than arguments) that are generated in structurally high positions. The ordering of these clitics with respect to each other, however, has been assumed until now to be due to the following template applying at PF or in a post-syntactic morphological component of the grammar:

(259)

	1	*2*	*3*	*4*
	kho	ba	am	am
				mo
	'indeed'	'will'	1PL, 2PL	1PL, 2PL
	5	*6*	*7*	*8*
	mee	dee	yee	no
	1SG	2SG;	3SG, 3PL	'then'
		'should'		

It is tempting to wonder whether the order of morphemes within the clitic cluster might reflect their syntactic derivation, in the same way that the structure of morphologically complex words has been hypothesized to reflect their syntactic derivation. Such a principle has been stated informally as follows (Baker 1985: 375):

(260)   The Mirror Principle

Morphological derivations must directly reflect syntactic derivations (and vice versa).

This chapter shows that a good portion of the template in (259) may also be derived by the syntax. Section 4.2 examines the pronominal clitics (slots 3–7), showing that they are not merely

ordered (roughly) by a ranking of first-person > second-person > third-person, but rather that the clitics are actually interpreted in those positions—a syntactic effect. Section 4.3 integrates the modal clitics (from slots 2 and 6) into syntactic structure, and section 4.4 shows that the adverbial clitics from the periphery of the cluster (slots 1 and 8) are amenable to a similar treatment. Finally, section 4.5 discusses non-syntactic aspects of clitic placement; as there is no syntactic rule of 'clitic movement', clitics may dislocate from their base position only by Prosodic Inversion (Halpern 1995), a last resort operation that applies in sentences containing, besides the clitic(s), only a verb. This section also examines how the overall analysis of clitics deals with some much-discussed facts of Pashto, showing that the mainly syntactic analysis of clitics as agreement morphemes explains many problems that hitherto have been considered the domain of phonology.

4.2. Pronominals

Possessive clitics offer striking evidence that pronominal clitics are ordered by the syntax. Chapter 3 showed, first, that when the argument structure of a transitive verb is not saturated by overt NPs in a sentence that has one overt NP and two pronominal clitics, the clitics are ambiguous in interpretation, because either of the clitics may saturate the argument structure, leaving the remaining clitic to be interpreted as the possessor of the single overt NP (sections 3.5.1 and 3.5.3). This ambiguity was due to the fixed ordering of the clitics with respect to each other, and the freedom of the empty pronominals to move to the specifier of either clitic head. At first glance, this rigid ordering of clitics would appear to be morphological in nature, due to their ambiguous interpretations. However, this freedom of clitics to be interpreted in any position was shown in section 3.5.4 to be absent in transitive sentences having two overt NPs and a single clitic: as the two overt NPs saturated the argument structure of the verb, the pronominal clitic had to be interpreted as a possessor of only the higher overt NP (not, ambiguously, either the higher or lower NP).

This ability to force a pronominal clitic to take on a possessive function will now be applied to see what it reveals about the ordering of clitics with respect to each other, by putting

two pronominal clitics in a transitive sentence having two overt NPs. Because the two overt NPs saturate the verb's argument structure (i.e., there is no *pro*), neither of the two clitics has the option to identify the subject or object, and so they can only be interpreted as genitive. In such a sentence, given what we have seen until now, several things might be expected to happen. Given that a single (possessive) clitic was shown to associate with the highest available NP, it might be expected that possessor associations with lower NPs would be ungrammatical. On the other hand, ambiguity might be expected, with a free choice for each possessive clitic as to which overt NP it associates with. Neither of the above options happens, however. Rather, the order of clitics within the cluster becomes important for interpretation in such cases, and the clitics exhibit a locality effect, each associating with its nearest NP. This is demonstrated in (261–262) for the clitics 1sg *mee* and 2sg *dee*, with different combinations of tense (past and present) and word order (SOV and OSV).

(261)　　a.　khezee　　mee　dee　kitaab　　lwest-e

　　　　　　　wife(OBL)　*1SG*　*2SG*　*book(DIR)*　*read(PAST)-MASC3SG*

　　　　　　　(i)　　'my wife was reading your book'

　　　　　　　(ii)　*'your wife was reading my book'

　　　　　b.　kitaab　　mee　dee　khezee　　lwest-e

　　　　　　　book(DIR)　*1SG*　*2SG*　*wife(OBL)*　*read(PAST)-MASC3SG*

　　　　　　　(i)　　'your wife was reading my book'

　　　　　　　(ii)　*'my wife was reading your book'

(262)　　a.　kheza　　mee　dee　kitaab　　lwel-i

　　　　　　　wife(DIR)　*1SG*　*2SG*　*book(DIR)*　*read-PRES3SG*

　　　　　　　(i)　　'my wife is reading your book'

　　　　　　　(ii)　*'your wife is reading my book'

181

b.	kitaab	<u>mee</u>	<u>dee</u>	kheza	lwel-i

  book(DIR)	*1SG*	*2SG*	*wife(DIR)*	*read-PRES3SG*

(i)	'your wife is reading my book'

(ii)	*'my wife is reading your book'

<div align="right">(Jan Mohammad, p.c.)</div>

Although the present tense sentences in (262) require both NPs to bear direct case (it is only in past-tense that ergative case-marking appears on the subject, resulting in a free word order), both SOV and OSV word orders are available, because the entities denoted by these NPs, 'wife' and 'book', are felicitously regarded as agent and patient respectively, regardless of their surface position (cf. Babrakzai 1999: 61). The relevant point of (261–262), then, is that scrambling (an operation in overt syntax) feeds the interpretation of the clitics. The same point is demonstrated below for 2sg *dee* and 3sg *yee*:

(263)	a.	spi	<u>dee</u>	<u>yee</u>	pisho	khog	krr-a

    dog(OBL)	*2SG*	*3SG*	*cat(DIR FEM)*	*hurt*	*do(PAST PERF)-FEM3SG*

(i)	'your dog hurt his cat'

(ii)	*'his dog hurt your cat'

  b.	pisho	<u>dee</u>	<u>yee</u>	spi	khog	krr-a

    cat(DIR FEM)	*2SG*	*3SG*	*dog(OBL)*	*hurt*	*do(PAST PERF)-FEM3SG*

(i)	'his dog hurt your cat'

(ii)	*'your dog hurt his cat'

(264)   a.   spay   <u>dee</u>   <u>yee</u>   pisho   khog-aw-i

dog(DIR)   *2SG*   *3SG*   *cat(DIR)*   *hurt-TRANS(PRES IMPF)-3SG*

(i)   'your dog is hurting his cat'

(ii)   *'his dog is hurting your cat'

b.   spay,   pisho   <u>dee</u>   <u>yee</u>   khog-aw-i[67]

dog(DIR)   *cat(DIR)*   *2SG*   *3SG*   *hurt-TRANS(PRES IMPF)-3SG*

(i)   'the dog, your cat is hurting him'

(ii)   *'your cat is hurting his dog'

(iii)   *'his cat is hurting your dog'

(iv)   *'your dog, it is hurting his cat'

(v)   *'his dog, it is hurting his cat'

(Jan Mohammad, p.c.)

The sentences of (261–264) show that when the verb's argument structure is entirely satisfied by full NPs, clitics associate left-to-right, as possessors, with the highest available NPs. This fact reveals that the order of clitics with respect to each other is not determined by a morphological component, but rather that it is the syntax that determines that first-person clitics are generated higher than second-person clitics, which in turn are generated higher than third-person clitics. If the clitics were ordered $1 > 2 > 3$ in the morphological component or at PF, the fixed interpretations demonstrated above would be entirely unexpected. Because this interpretive effect is caused by the order of the clitics themselves, and because overt syntax (not morphological

---

[67] Because of the reciprocal participants ('dog' and 'cat'), the OSV interpretation is available in present tense only with a heavy pause (indicated by a comma) after the dislocated object, requiring the 3sg clitic *yee* to be interpreted as resumptive, rather than possessive. Note that this fact shows that possessive interpretation of the clitic is restricted to its clause (TP), and that the clitic may not associate with an NP outside of its clause—even if only topicalized.

structure or PF) feeds interpretation (LF), the clitics have been merged in their surface positions directly by the syntax, without any need for internal ordering by a morphological template.

In many languages, there is a morphosyntactic split between first- and second-person, on the one hand, and third-person on the other. In K'ichee', for example, when a first- or second-person argument competes with a third-person argument for agreement, it is always the first- or second-person argument—not the third—that triggers agreement (Hale and Storto, n.d.). This person-split is also found in most Salish languages, which evince a pattern of split-ergativity in pronominal inflection: first- and second-persons are inflected on a nominative/accusative pattern, while third-person is inflected on an ergative/absolutive pattern (Czaykowska-Higgins and Kinkade 1997: 32–33). In Basque, first- and second-persons show nominative/accusative agreement on the verb, while third-person arguments show ergative agreement (Fernández 1999: 181). Still other examples of such person-splits are mentioned by Aissen (1997: 707–708). It was seen in chapter 1 that Pashto, too, treats first-and second-person together in the cases that are assigned to strong pronouns; see the paradigm in (21). Here, then, is another area of grammar in which first- and second-person are treated together, in opposition to third-person, as may be seen from the fact that the 1/2pl clitics *am* and *mo* precede 3sg/pl *yee*, although the singular clitics evince an even tighter ranking, with 1sg > 2sg > 3sg. It has never been clear whether person-splits are due to some property of discourse (since first- and second-person referents are discourse participants, whereas third-person referents are not), and even less clear how to account for these splits formally (whether as a ranking of persons, or as syntactic structures). The interpretation of clitics in Pashto, though, strongly suggests that clitics are actually merged, according to their person, directly into the sentence, and are not subject to reordering after Spell Out.

4.3. Modals

The preceding material has dealt exclusively with the pronominal 2P clitics. It remains to integrate the non-pronominal 2P clitics into this analysis. Ignoring the modal clitic *dee* for the

moment, a glance at the template in (259) shows that the two kinds of non-pronominal clitics—the adverbials and modals—appear at the periphery of the cluster. A templatic approach to their ordering gives the impression that it is accidental. However, these further aspects of clitic ordering may be derived by the syntax, and it is surely desirable to do so, insofar as the relevant syntactic principles are independently required. For example, the fact that the modal *ba* 'will' occurs in column 2 in the template of (259), preceding the pronominal clitics, reflects the fact that modals are generated in a position higher than agreement, either under the Tense node or in distinct Modal heads (cf. Chomsky 1981: 140 n. 28; Cinque 1999: 78 ff.). Consider the following sentence, in which switching the order of the 2P clitics *ba* and *dee* would result in ungrammaticality:

(265)　　　Tor　　ba　　dee　　wu-win-i[68]

　　　　　Tor　　will　　2SG　　PERF-see-3SG

　　　　　'Tor will see you' (Tegey 1977: 187)

If the structure of this sentence is as follows, with the subject *Tor* having moved to Spec/TP due to the EPP, the syntax alone has derived the second position of the clitics, as well as their internal order:

(266)　　　[TP　Tor　[ModP　ba　[CliticP　dee　[AspP　wu-win-i]]]]

　　　　　　　Tor　　　　　will　　　　　2SG　　　　　PERF-see-3SG

Evidence for generating modal clitics in this position is that there are strong forms of modals that appear in the same position. An example is *baayad* 'should', which, not coincidentally, appears in 'second position' in the following sentences:[69]

---

[68] The vowel of perfective *we* is labialized here due to the following labial consonant.

(267)   a.   asad   **baayad**   yawa   baja   raas-i

                Asad   *should*   *one*   *hour*   *come(PRES PERF)-3SG*

                'Asad should be here at one'

    b.   te   **baayad**   pe   dee   po   see

                PN2SG   *should*   *about*   *this*   *informed*   *become(PRES PERF2SG)*

                'you should know this'

    c.   laylaa   **baayad**   kor   ta   laarr-a   ne   si

                Layla   *should*   *house*   *to*   *way-FEM3SG*   *NEG*   *become(PRES PERF3SG)*

                'Layla shouldn't go home' (Tegey and Robson 1996: 149)

When the full NP subject is omitted, however (being able to be identified by the verbal agreement), the strong modal becomes initial, which a 2P clitic would never be able to do:

(268)   a.   **baayad**   yawa   baja   raas-i

                should   *one*   *hour*   *come(PRES PERF)-3SG*

                'he/she should be here at one'

    b.   **baayad**   pe   dee   po   see

                should   *about*   *this*   *informed*   *become(PRES PERF2SG)*

                'you should know this'

---

69 Strong (non-clitic) modals like *baayad* 'should' have been borrowed from Persian (Babrakzai 1999: 58), and are used only in some dialects, and in educated speech (Jan Mohammad, p.c.).

c. **baayad**   kor   ta   laarra   ne   si

should   *house*   *to*   *way-FEM3SG*   *NEG*   *become(PRES PERF3SG)*

'he/she shouldn't go home' (Jan Mohammad, p.c.)

Because of the possibility of scrambling, the strong modal may also follow constituents that have been fronted. Compare (268c) to the variants below:

(269)   a.   kor   ta   **baayad**   laarra   ne   si

house   *to*   *should*   *way-FEM3SG*   *NEG*   *become(PRES PERF3SG)*

'he/she shouldn't go home'

b.   kor   ta   laarra   **baayad**   ne   si

house   *to*   *way-FEM3SG*   *should*   *NEG*   *become(PRES PERF3SG)*

'he/she shouldn't go home' (Jan Mohammad, p.c.)

On the other hand, the strong modal may not occur so far to the right that it would follow negation—a sure sign that the modal is generated in a position higher than Neg:

(270)   *kor   ta   laarra   ne   **baayad**   si

house   *to*   *way-FEM3SG*   *NEG*   *should*   *become(PRES PERF3SG)*

'he/she shouldn't go home'

This behavior of the strong modal *baayad* 'should' contrasts strikingly with its weak, 2P clitic counterpart. As was illustrated in (84c,d) in the previous chapter—repeated below as (271a,b)—the 2P modal clitic *dee* 'should' may appear to the right of negation if it would otherwise lack a phonological host to its left. The modal clitic *ba* must likewise follow negation if it would otherwise lack a phonological host to its left, as shown in (272):

(271)   a.   khar   <u>dee</u>   ne   raawal-i

              donkey   should   *NEG*   *bring-PRES3SG*

              'He should not bring the donkey'

         b.   ne   <u>dee</u>   raawal-i

              NEG   should   bring-PRES3SG

              'He should not bring it' (Tegey 1977: 82–83)

(272)   ne   <u>ba</u>   <u>dee</u>   pezan-i

         NEG   may   2SG   know-PRES3SG

         'Maybe he doesn't know you' (Tegey 1977: 84)

The simplest explanation of this contrast is to assume that all modals are generated in the same, structurally high position—ModalP—where they remain throughout the derivation. The difference between the strong modal *baayad* 'should' and its weak counterpart *dee* 'should' is that the latter is prosodically categorized for a host to its left—a typical requirement for a clitic. Movement is allowed only for the modal clitic, if remaining in its base position would result in its not having a phonological host to its left, as has happened in creating the contrast illustrated in (271) above. Such movement occurs only at PF, as a last resort strategy to save a syntactically well-formed (but phonologically ill-formed) structure. Further such examples will be discussed in section 4.5.

4.4. Adverbials

The remaining class of items, the adverbials, are not 2P clitics in all varieties of Pashto. Turning first to Tegey (1977), for whom they are indeed 2P clitics, the adverbials are also amenable to an analysis in which their position reflects their base (syntactic) order. That the adverbial 2P clitics

kho 'indeed; really; of course' and *no* 'then' appear on the extreme left and right (respectively) periphery of the clause, for example, is reminiscent of a recent idea that adverbials may appear in the specifier of various functional projections, such as Mood, Tense, Aspect, and Voice (Cinque 1999), but would not be expected to be interleaved among agreement projections, disrupting their homogenous character to produce a sequence of clitics such as schematized in (273), which would in fact be ungrammatical:

(273)   *1SG ADV 3SG

Cinque's (1999: ch. 5) treatment of adverbs assumes that they appear in fixed positions (due to their being in specifier positions of functional projections, which likewise appear in a rigid order), and so when DPs are interleaved among adverbs, it is because the DPs are appearing in AgrPs (what Cinque calls 'DP-related functional projections'), which may appear between the adverbial functional projections, but which do not themselves host adverbs in their specifiers. The position of adverbs on the left and right periphery of the clause therefore suggests that the Pashto clause has functional projections that host adverbs, with intervening Agr phrases (which are headed by the 2P pronominal clitics). Ignoring the difficulty of locating the specific adverbs in one projection versus another, we might suppose that the rightmost adverbial *no* 'then' is in Spec/AspP (since we already have ample evidence for the existence of that projection and its position above VP), and that the leftmost adverbial *kho* 'indeed' occupies Spec/ModalP, elaborating the representation in (266) thus:

(274)

```
                    TP
                   /  \
                ModalP
                /    \
            kho      Modal'
          'indeed'   /    \
                  Modal   CliticP
                   |      /    \
                  ba   Clitic  CliticP
                 'will'  |     /    \
                       mee  Clitic  AspP
                       1SG    |     /   \
                            yee    no    Asp'
                            3SG  'then'  /\
                                        /__\
                                      VP . . .
```
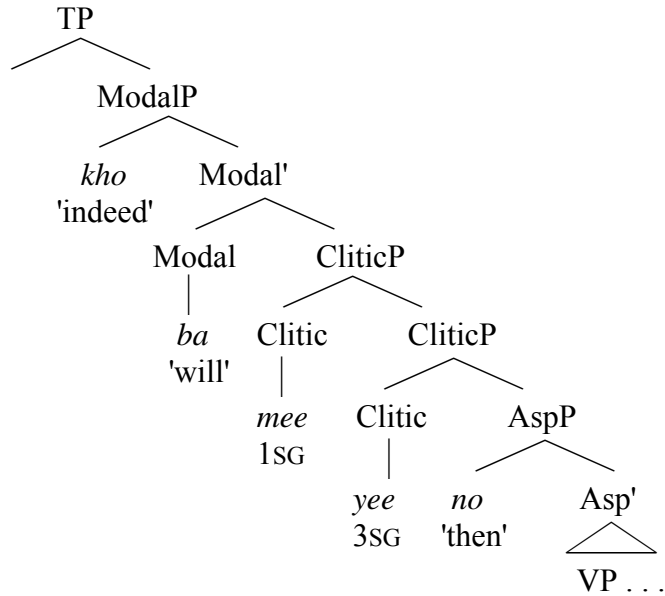
This structure offers a striking advantage over the template in (259), since the order of clitics with respect to each other is derived by independent principles. This structure also explains the 'second position' effect displayed by these clitics, about which a template has nothing to say: due to the EPP, and/or the option of scrambling to a clause-initial topic/focus site, Spec/TP (or a higher topic/focus position) will usually be occupied by an overt constituent that has moved from its VP-internal base position. In such cases, the clitics illustrated in (274) will have their requirement for a leftward prosodic host vacuously satisfied; when no such movement to Spec/TP occurs, however, a last resort PF option of Prosodic Inversion applies (Halpern 1995), which will be illustrated in the following section. The notion of 'second position' is thereby reduced to a descriptive artifact of the effects of independent structures and processes.

In other varieties of Pashto, the adverbial clitics are not actually second-position clitics. Mohammad (1993: n. 2) believes that *kho* 'indeed' is a focus marker, rather than a clitic, and Babrakzai (1999: 47) also remarks that their distribution is different from the pronominal clitics. In (275a), the adverbial 'clitic' *no* 'then' appears initially (which a pronominal clitic could never do), while (275b) is particularly striking in showing the adverbial 'clitic' *kho* 'indeed' separated from the pronominal 2sg clitic *dee*, which appears in canonical second position:

(275) a. **no**   ze   de   feelem   na   wrusta berta   raaghl-em

           then   *PN1SG* *of* *movie* *from* *after* *return* *came-1SG*

           'Then I came back after the movie'

       b. wror   ta   <u>dee</u>   teeleepun   **kho**   we-krr-a

           brother *to*   *2SG*   *telephone*   *indeed*   *PERF-do-IMP*

           'You (should) indeed call up your brother!' (Babrakzai 1999: 47)

A sentence like (275b) provides a notable contrast to Tegey's (276), in which *kho* 'indeed' appears to be a 2P clitic, appearing in second position with the pronominal clitics. And the sentence in (312c) below, from Tegey's own work, constitutes further evidence for treating *kho* separately from the other clitics.

(276) agha   dzhega   pezrreporee   sra   maanney

       that   *tall*   *interesting*   *red*   *building*

       <u>kho</u>   <u>ba</u>   <u>dee</u>   khwakha   wi

       indeed   *would*   *2SG*   *liked*   *were*

       'You would have indeed liked that tall, interesting red building' (Tegey 1977: 83)

Aside from this variation, however, it is clear that adverbial clitics do not intervene among the pronominal clitics. The explanation for this fact is that adverbs are merged as the specifiers of functional projections, and remain in those positions throughout the derivation. For Tegey's variety of Pashto, the adverbial clitic *kho* 'indeed' is prosodically subcategorized for a host to its left. Varieties of Pashto in which adverbial 'clitics' do not pattern with the other pronominal clitics—exemplified by (275)—can only be taken as further evidence for an analysis in which

adverbials appear as specifiers of functional projections; they differ from Tegey's Pashto in not requiring a host to their left, hence they may appear initially, as in (275a).

Crucially, however, the adverbials do not intervene among the pronominal clitics, which occupy the heads of agreement projections; in other words, a pattern like (273) above does not obtain. This fact is not entirely predicted by Cinque's (1999) account of adverb placement, if only because he does not discuss such an asymmetry. Since his system permits DP-related functional projections (AgrPs) to intervene among the adverbial positions, it might be expected that (273) would be a possible pattern. That it is *not* a possible pattern lends support to the previous chapter's main proposal that clitics occupy the heads of agreement phrases that are freely generated between TP and AspP, and which bear oblique case features: because of the ambiguity of the clitic pronominals with respect to their function (as ergative, accusative, or genitive)—due to their each being related to a single case (oblique)—it is not possible to scatter the CliticP projection throughout the structure, in the same way that an agreement projection specifically for objects could conceivably appear between other functional projections that hosted adverbs in their specifiers.

We will return to some remaining issues of clitic ordering in section 4.5.3.

4.5. Morphophonological aspects of clitic ordering
4.5.1. Prosodic inversion

Having demonstrated that most of the placement and interpretation of second-position clitics may be derived from their remaining in the syntactic positions in which they are merged, a final class of examples remains to be explained. These examples have attracted the most attention in studies of 2P clitics, for the reason that it is unusual for a second-position clitic to intervene among the parts of a verb. The most productive example of this behavior involves a clitic intervening between the perfective morpheme *we* and a verb stem.

The imperfective aspect of a monomorphemic verb is illustrated in (277a). As was discussed in chapter 2, perfective aspect is formed by adding the perfective morpheme *we* to the verb, which attracts stress from the root, as shown in (277b).

(277)   a.   tor   sra   skund-él-a

                  Tor   *Sra*   *pinch-PAST-FEM3SG*

                  'Tor was pinching Sra'

           b.   tor   sra   wé   skund-el-a

                  Tor   *Sra*   *PERF*   *pinch-PAST-FEM3SG*

                  'Tor pinched Sra' (Tegey 1977: 85)

When occurring with imperfective monomorphemic verbs, the clitic follows the verb, as illustrated below:

(278)   a.   matsh-aw-él-ee   <u>yee</u>

                  kiss-TRANS-PAST-2SG   *3SG*

                  'He was kissing you'

           b.   tekhn-aw-él-a   <u>mee</u>

                  tickle-TRANS-PAST-FEM3SG   *1SG*

                  'I was tickling her' (Tegey 1977: 86)

Because VP, like other lexical categories, is head-final, all sentences in which the clitic follows the verb are claimed here to involve prosodic inversion (Halpern 1995), a PF operation that applies as a last resort, minimally moving the clitic rightward until it finds a prosodic host to its

left. Taking (278a) as an illustration, the syntax derives the following structure at LF, which is also representative of the Spell Out representation (since nothing overt has moved):

(279)   [$_{TP}$ ... [$_{CliticP}$ pro$_1$ [$_{Clitic'}$ yee   [$_{VP}$ pro$_2$ matsh-aw-él-ee]]]]

                                         3SG   *kiss-TRANS-PAST-2SG*

This structure is entirely licit at LF: the sentence is past-tense (i.e., ergative), and so the object, *pro$_2$*, is identified by the 2sg agreement suffix on the verb. The subject, *pro$_1$*, has moved from its base position in Spec/*v*P (not shown above) to Spec/CliticP, where it is identified by the 3sg clitic *yee*, which heads the agreement projection. The highest overt lexical item is therefore the 3sg clitic *yee*, which has a prosodic subcategorization requiring a host to its left. At PF, the syntactic structure of (279) is erased, and word boundaries are derived from the positions of lexical heads. In the following PF representation, ω represents a phonological word, which comprises the contents of the syntactic head V.   (Recall that words enter the derivation fully inflected.) Although it occupies a syntactic head position, a word boundary is not derived for the 3sg clitic *yee*, as it is prosodically subcategorized for a host to its left. While it is a word for the syntax, it is not a word for the phonology:

(280)   yee [$_{ω}$ matshawélee]

Lacking any overt material to its left, *yee* minimally inverts, resulting in the following structure (assumed to be adjunction, though there are other possibilities that would best be explored in a phonological account of clitics):

(281)   [$_{ω}$ [$_{ω}$ matshawélee] yee]

At this point, other phonological rules apply. In this example, the clitic *yee* is phonologically reduced according to the rule formalized in (311) below, deriving the phonetic form:

(282)   [matshawéleey]

The LF and PF representations thus derived have different orders of the terminal elements, but this difference is irrelevant, as the structure must be well-formed at both levels in order for it to be grammatical. If the clitic remained in its base position at PF, the structure would be ill-formed at that level. Note also that, as in the syntactic derivations of 2P clitic placement in the previous chapter, there is similarly no 'second position' at PF, either. That clitics appear in second position is merely an artifact of minimal prosodic inversion, which could only derive a structure in which the clitic appears as the second element in a string of prosodically independent morphemes.

Evidence for the 'minimal' character of prosodic inversion comes from sentences containing only a polymorphemic verb and a second-position clitic. Recall from previous chapters such sentences as the following, in which a clitic may divide the morphemes of what has been (erroneously) regarded in some analyses as a single lexical item:

(283)   a.   wé   dee   ritt-e

                PERF   2SG   insult(PAST)-MASC3SG

                'You insulted him'

          b.   wú   ba   dee   gur-i[70]

                PERF   will   2SG   see-3SG

                'He will see you' (Tegey 1977: 87)

---

[70] The vowel of the perfective morpheme is rounded here because of the following consonant.

A contrast like the following shows that in sentences containing only a verb and perfective morpheme in addition to the clitics, the clitics must follow the perfective marker (apparently dividing it from its verb stem), but may not follow the verb:

(284)   a.   wáa   <u>dee</u>   khist-el[71]

                PERF   *2SG*   *buy-PAST(MASC PL)*

                'You bought them' (Laghman)

        b. *wáa   khist-el   <u>dee</u>

                PERF   *buy-PAST(MASC PL)*   *2SG*

        c. *<u>dee</u>   wáa   khist-el

                2SG   *PERF*   *buy-PAST(MASC PL)*

As was shown in chapter 2, the perfective morpheme *we* heads AspP, which is above the verb. Assuming for simplicity in representation that AspP is head-initial, the structure of (283b) at LF is as follows:

(285)   [$_{TP}$ ... [$_{ModP}$ ba   [$_{CliticP}$ pro$_{2[i]}$ [$_{Clitic'}$ dee

                        will   *2SG*

        [$_{AspP}$ we   [$_{vP}$ pro$_1$ [$_{VP}$ t$_i$ gur-i]]]]]]]]

                PERF   *see-3SG*

At LF, the structure is well-formed: the sentence is non-past-tense, and so *pro$_1$* is identified by the agreement suffix on the verb. The object, *pro$_2$*, moves to Spec/CliticP in order to be

---

[71] The appearance of perfective *we* as *waa* will be discussed in the next section.

identified. Its interpretation as the object derives from its base position as sister of the verb. Because only empty categories have moved, this representation also suffices to illustrate the Spell Out form. When PF strips away the syntactic information, inserting prosodic word boundaries at the edges of overt, non-clitic heads, the PF representation in (286) is derived.

(286)   ba dee [$_\omega$ we] [$_\omega$ guri]

Perfective *we* is both a syntactic head and a proclitic; because it has a host to its right, it may form its own phonological word at this stage of the derivation. While they are also words in the syntactic sense, the enclitics *ba* and *dee* need a host, and so they do not form prosodic words on their own; both of them will need to invert together in order to find a host to their left. The fact that they invert together is what gives the appearance of a single 'cluster' of clitics. In previous sections, the 'cluster' effect (i.e., that the clitics follow each other in the order that they do) was shown to be largely derivable from their underlying syntactic representation, and the same cluster effect at PF may similarly be derived.[72] It is sufficient for the purposes here to assume that the prosodic structure is derived as fully as possible; there is not yet a need for prosodic inversion, because the enclitic *dee* actually has a prosodic host to its left: the enclitic *ba*. The clitic *dee* may therefore prosodically adjoin to *ba*, deriving the following representation—but note that, crucially, *dee* cannot form a prosodic word with *ba*, since *ba* itself still needs a host to its left:

(287)   [[ba] dee] [$_\omega$ we] [$_\omega$ guri]

It is the prosodic requirement of the leftmost clitic (the head of this constituent)—that it needs a host to its left—that compels the entire category thus derived to invert: it is the last resort to save

---

[72] Many technical, phonological questions arise at this point, but they will not be explored; consult Halpern (1995) for discussion and references.

an otherwise phonologically illicit structure. Because prosodic inversion is minimal, this constituent inverts with the prosodic word to its immediate right (perfective *we*), adjoining to it:

(288)   [$_\omega$ [$_\omega$ we] [[ba] dee]] [$_\omega$ guri]

The representation in (288) is the surface order that was illustrated by (283b), and the schwa of perfective *we* will be subject to a further rule of labialization, as mentioned in the footnote to (283b); see (265) above and (305b) below for other examples of labialization. Note again, though, that there is no 'second position' for the clitics that has any more than accidental status; it is the effect of minimal prosodic inversion that merely appears to place clitics in a special second position.

It has often been supposed that in such sentences, the clitic is actually intervening among the parts of the verb, making the language appear typologically unusual for a reason that is not correct. Although Tegey regards perfective *we* as a prefix, he does not consistently transcribe it as such—which is understandable when one considers that it does not appear to be a prefix when clitics follow it. As was noted in chapter 2, this morpheme is best regarded as occupying a separate syntactic projection (the head of AspP), and hence it forms its own phonological word at PF. It is exactly the behavior of this morpheme with respect to clitics that demands such a treatment. In contrast, verbal suffixes may never be divided separated from the verb stem by clitics, as was illustrated by (33) in chapter 1. If perfective *we* is an affix, it is clearly not as tightly bound to the root as are the agreement suffixes; the perfective morpheme is only loosely affixal, and so it is best treated as a proclitic, as was suggested in chapter 2.

There are other morphemes, though, that are more clearly prefixes on the verb, and which are not as likely as perfective *we* to head their own maximal projections. For a subset of these verbs, their imperfective forms may have stress either near the end of the verb (the usual pattern) or initially. In the latter case, the clitic will intervene between the prefix and the verb stem. The contrast is exemplified in (289–290).

(289)   a.   a-khistéle   <u>mee</u>

buy   *1SG*

'I was buying them'

b.   á   <u>mee</u>   khistele

PREFIX   *1SG*   *buy*

'I was buying them'

(290)   a.   a-ghuste   <u>mee</u>

wear   *1SG*

'I was wearing it'

b.   á   <u>mee</u>   ghuste

PREFIX   *1SG*   *wear*

'I was wearing it' (Tegey 1977: 89)

Crucially, in monomorphemic verbs that show the same stress alternation, clitics may not divide the verb, but rather must follow it:

(291)   a.   saatem   <u>yee</u>

keep   *3SG*

'I keep it'

b.   sáatem   <u>yee</u>

keep   *3SG*

'I keep it'

(292)   a.   pereebde   <u>mee</u>

              beat   *1SG*

              'I was beating him'

    b.   péreebde   <u>mee</u>

              beat   *1SG*

              'I was beating him' (Tegey 1977: 88)

These paradigms confirm that morpheme structure (rather than stress alone) contributes to the determination of 'second position'.

The alternation in (289–290) also applies to compound verbs, as was discussed in chapter 2. Here, the perfective is formed not by adding the perfective morpheme *we*, but rather by shifting stress to the initial syllable. Regardless of the source of the stress shift (optionality versus perfective formation), however, initial stress on a verb prefix licenses the prefix to host clitics. The imperfective versus perfective alternation and its interaction with clitic placement is exemplified below:

(293)   a.   tteel-waahé   <u>mee</u>

              push   *1SG*

              'I was pushing it'

    b.   ttéel   <u>mee</u>   waahe

              PREFIX   *1SG*   *push*

              'I pushed it' (Tegey 1977: 92)

(294)   a.   ttak-waahé   <u>mee</u>

                   shake   *1SG*

                   'I was shaking it'

        b.   tták   <u>mee</u>   waahe

                   PREFIX   *1SG*   *shake*

                   'I shook it' (Tegey 1977: 92)

(295)   a.   poree-weesté   <u>mee</u>

                   carry   *1SG*

                   'I was carrying it across'

        b.   póree   <u>mee</u>   weeste

                   PREFIX   *1SG*   *carry*

                   'I carried it across' (Tegey 1977: 92)

The data of (289–295) support the suggestion of chapter 2: perfective aspect is a strong feature. Its appearance in the Asp node, directly above VP, compels the nearer part of the verb to move there. The verbs involved may be compound or bearing a prefix. It is worth noting, though, that these latter verbs not contain free morphemes, as do the compound verbs discussed in chapter 2, which are productively formed by combining adjectives and nouns with transitive and intransitive auxiliaries. The *a*-initial verbs exemplified in (289–290), especially, are the only vowel-initial verbs in the language, and constitute a rather small class:

(296)   akhistel   'to buy, to take'

         aleyel   'to singe, to roast'

         atshawel   'to throw'

         aghustel   'to put on, to wear'

         alutel, alwézem   'to fly'

         ákhssel   'to knead'

         árwem, áwrem   'to hear'

         áwrrem   'to turn'

         astawem   'to send'

Other vowel-initial words identified by Morgenstierne (1927) are listed below, although his initial rounded vowels are more likely glides:[73]

(297)   aazziyil   'to incite, to stimulate'

         annel   'to grind'

         akheerrel   'to plaster, to besmear'

         udel, uwem   'to weave'

         orbal   'to curl'

         oreedel   'to rain'

         óseedel   'to dwell'

Darmesteter (1888–90: cxxxix) remarks that the initial *a* is the same prefix as in Avestan, which had a directional flavor. While Morgenstierne (1927) reconstructs a prefix for some of these verbs, he never states explicitly that its source was indeed Avestan. Although it is plausible that

---

[73] See Bell and Saka (1983) for discussion of these sorts of initial consonant clusters.

initial *a* was indeed a prefix in the prehistory of Pashto, it is clearly no longer so.[74]   In this

respect, the bound morphemes in such verbs are similar to English words like the following,

whose morphological constituency is somewhat opaque (as has been noted by Kaisse 1981: 203–

204, also in the context of Pashto clitic placement):

(298)   permit, remit, transmit, commit, compel, concur, recur, transfer

In English, the analysis of such words as those in (298) into bound prefixes and stems (e.g., *per-*,

trans-, *con-*, *-mit*, *-pel*, etc.) is internal to the lexicon, without syntactic consequences (Chomsky

and Halle 1968: 94). And yet even in English, some prefixes are separable from their stems (by *f*-

word infixation, for example), and may even be conjoined:

(299)   a.   pre- and post-millennium festivities

            b.   pro- and anti-abortion forces

Although similar bound morphemes in Pashto appear to induce syntactic effects (in that they

affect clitic placement), such a conclusion would be premature, and based on the assumption that

clitic placement is syntactic. As the previous chapter has argued, there is no syntactic rule of

clitic placement; the seemingly exceptional placement of clitics in such forms must be regarded as

the simple and familiar PF operation of prosodic inversion. This conclusion is expected, as there

is no reason to expect that the correct treatment of these bound morphemes in Pashto would need

---

[74] Most of the Pashto speakers I have consulted do not allow *a*-initial verbs to be divided by
clitics (contra Tegey 1977), and of the small number who do, they do so with only some of the
verbs, and not others. Verbs requiring that perfective aspect be formed by shifting stress to the
initial syllable are also not especially productive among my consultants, most of whom prefer to
mark perfective aspect with the more productive *we* proclitic; 2P clitics may indeed intervene
between this proclitic and the verb stem, suggesting that the perfective 'prefix' occupies a separate
syntactic projection—the analysis that has already been advanced.

to be appreciably different from that accorded to similar morphemes in English, given how strongly constrained language is hypothesized to be.

Further evidence for this analysis is that absence of a stressed host near the beginning of a sentence compels a clitic to appear further to the right of the syntactic second position. If the first constituent of a sentence does not bear at least one main stress, the clitic will occur further to the right, following the first unit that does bear stress. This point is illustrated by the following data:

(300) a. [PP pe rasey] <u>ba</u> <u>yee</u> wé-tarri

      with rope will 3SG PERF-tie

     'He will tie it with the rope'

    b. [PP pee] wú <u>ba</u> <u>yee</u> tarri

      with-it PERF will 3SG tie

     'He will tie it with it'

    c. [PP pee] tarri <u>ba</u> <u>yee</u>

      with-it tie will 3SG

     'He will be tying it with it'

(301) a. [PP laylaa na] <u>dee</u> aa-khiste

      Layla from 2SG PREFIX-buy

     'You were buying it from Layla'

    b. [PP tree] á   <u>dee</u> khiste

      from-her PREFIX 2SG buy

     'You were buying it from her'

c. [PP tree] aa-khiste dee

      from-her   PREFIX-*buy*   2SG

'You were buying it from her' (Tegey 1977: 114)

The (a) sentences of (300–301) have a full pre/postpositional phrase in initial position, and—because the nominal complements in such PPs constitute prosodic words in their own right—the clitic appears after that first, stressed constituent. Two structures are possible: the PP may have been scrambled to (or merged in) a clause-initial topic/focus site, allowing the clitics to remain *in situ*, or else the clitics have undergone prosodic inversion, and must invert with the PP before they can find a phonological host. In contrast, the initial PP of the (b) and (c) sentences comprises the proforms *pee* 'with it' and *tree* 'from her'. Because these proforms are always stressless (and may even be proclitic themselves), the 2P clitic may not be hosted by them. Rather, in the (b) sentences, a preverbal morpheme that bears stress hosts the clitic. In the (c) sentences, only the verb root itself bears stress, and so the clitic may have an appropriate host only by appearing in final position. That prosodic inversion ignores initial, stressless constituents, is even more strikingly illustrated by such sentences as the following, in which several stressless constituents may appear initially and are unable to host the clitic:

(302)   a. [PP ra ta] [PP tee] [AdvP raa] shkaawe dee

           me for   *from-it*   *here*   *pick*   2SG

        'You were picking it for me from it (and bringing it) here'

   b. [PP ra ta] [PP tee] [AdvP raa] we dee shkaawe

           me for   *from-it*   *here*   PERF   2SG   *pick*

        'You picked it for me from it (and brought it) here' (Tegey 1977: 119)

As shown above, the clitic behaves as if the initial, stressless constituents were absent, but otherwise is positioned with respect to stress, morpheme structure, and constituency, as in the usual cases.

4.5.2. Vowel coalescence

The relevant sentences of (94) and (284) above, repeated below as (303a,b), illustrate Pashto's notorious 'vowel coalescence', first discussed by Tegey (1977). Note that the perfective morpheme, which is usually *we*, becomes *waa* when it is attached to the verb root *akhistel* 'to buy', one of the *a*-initial verbs of (296), whose initial vowel has already been seen in (289) to be separable from the stem. The vowel coalescence rule is formalized in (304). In (303a), the perfective marker precedes the verb stem, and so the environment for vowel coalescence— adjacency—is met. Strikingly, however, in (303b), the perfective marker is separated from the verb stem by the 2P clitic *dee*, and yet vowel coalescence still occurs:

(303)   a.   parun   <u>dee</u>   wáa   khist-el

         yesterday   *2SG*   *PERF*   *buy-PAST(MASC PL)*

         'You bought them yesterday'

         b.   wáa   <u>dee</u>   khist-el

               PERF   *2SG*   *buy-PAST(MASC PL)*

               'You bought them'

(304)   Vowel Coalescence (Kaisse 1981: 202)

             [e]$_{particle}$ + [a, aa]$_{verb}$ → [aa]

Because he assumed that clitic placement was a syntactic process, sentences like (303b) led Tegey (1977) to suppose that a syntactic rule (clitic placement) needed to occur after a

phonological rule (vowel coalescence). This conclusion was troubling for a strictly derivational theory of grammar in which syntactic rules apply before phonological ones.

As was discussed above, Kaisse (1981) challenged Tegey's conclusion by regarding a class of seemingly monomorphemic verbs as polymorphemic, although she retained Tegey's assumption that clitic placement was a syntactic process. Under the proposal here that clitics are never moved in the syntax, but rather may only move in the phonology as a last resort to find a host to their left, the troubling case of vowel coalescence in (303b) is reduced to an ordinary question about rule ordering in the phonology—exactly the component of the grammar in which rule ordering obtains (Bromberger and Halle 1989). The 2sg clitic *dee* in (303b) has been generated as the head of CliticP (an agreement projection), which is higher than AspP and VP, as shown by its position with respect to perfective *we* in (303a). Vowel coalescence between the perfective morpheme and the initial vowel of the verb stem occurs at this point—crucially, before prosodic inversion applies.

The fact that the vowel coalescence applies before prosodic inversion constitutes additional evidence for the 'last resort' nature of this operation. Ordinary phonological rules proceed as usual, and it is only when all of these rules have applied, and the clitic still lacks a host to its left, that prosodic inversion must apply, inverting the clitic with the perfective prefix. It is also no longer surprising that the initial vowel of the verb stem is omitted after prosodic inversion: it was already shown in the previous section that this is a historic prefix, independently separable from the verb stem; because this vowel has coalesced with the vowel of the perfective morpheme *we*, deriving *waa*, there would be no way for the clitic to invert any more minimally than it already does (intervening among the phonemes of the perfective morpheme, for example, which it never does in ordinary cases of prosodic inversion), and so the clitic finds a minimal and sufficient host in perfective *waa*. Such examples are not puzzling in this account, as they are when one assumes that there exists a syntactic process of clitic placement.

4.5.3. Remaining issues of clitic ordering

Aside from the desirability of the syntax determining the order of clitics, as discussed throughout this chapter, there are two aspects of clitic ordering that may not be so easily derived by the syntax, but which rather appear to be morphophonologically determined. First, there is an interesting exception to a generalization like (273) with respect to the modal *dee* 'should', which might have been expected to appear in slot 2 of the template in (259)—or under the Modal node in (274)—along with the modal *ba* 'will' (Tegey 1977: 197). In fact, *dee* 'should' appears in column 6, with the homophonous 2sg *dee*. This accounts for the initially surprising fact that while Pashto may express phrases like 'I should' or 'he should' with 2P clitics, it may not express a phrase such as 'you (sg) should' with those same resources, but rather must resort to a strong pronoun for 2sg, as illustrated below for 'you (sg)' as an object. (Recall from the footnotes to (265) and (283b) that the vowel of the perfective morpheme *we* is labialized here, because it precedes a labial consonant.)

(305)   a.   *Tor   <u>dee</u>   <u>dee</u>   wu-wah-i

              Tor   *should*   *2SG*   *PERF-hit-3SG*

              'Tor should hit you'

      b.   Tor   <u>dee</u>   taa   wu-wah-i

              Tor   *should*   *PN2SG*   *PERF-hit-3SG*

              'Tor should hit you' (Tegey 1977: 196)

In fact, any sequence of identical clitics is excluded. In (306a), the 1sg sequence *mee mee* is ungrammatical; one of the 1sg pronominals must instead take the form of a strong pronoun, as in (306b):

(306)   a.   *wror   <u>mee</u>   <u>mee</u>   wah-i

              brother   *1SG*   *1SG*   *hit-3SG*

              'My brother is hitting me'

        b.   wror   <u>mee</u>   maa   wah-i

              brother   *1SG*   *PN1SG*   *hit-3SG*

              'My brother is hitting me' (Tegey 1977: 193)

Illicit clitic sequences such as the one in (306a) may be excluded straightforwardly by the template in (259), since one form is drawn twice from a single column—although such an account does not explain why clitics are put into those columns in the first place. As Tegey (1977: 194 ff.) notes, the explanation could not be that there is a morphosyntactic constraint on adjacent person features, since this would wrongly entail that (306b) should be as ungrammatical as (306a). Moreover, the contrast in (305) does not even involve adjacent pronominals; its ill-formedness is due to the fact that the adjacent clitics are phonologically identical.

        The ordering of clitics—and particularly the ordering of the modal clitic *dee* 'should' with respect to the pronominal clitics—therefore appears to be determined by a constraint like (307), familiar from phonology (Myers 1997 and references there):

(307)   Obligatory Contour Principle (OCP)

            Adjacent identical elements are prohibited.

Bearing in mind the difficulty of formulating this principle in morphological terms (Bonet 1995: 629; cf. Ross 1972), there is something akin to the OCP applying in two areas of the grammar: (i) at PF, within the clitic cluster itself, serving to block like sequences that are otherwise syntactically well-formed; and (ii) in some other component of the grammar, actually comparing members of the clitic paradigm and positioning them according to their shape, thereby placing

dee 'should' with respect to first- and third-person in the same position that the homophonous 2sg clitic occupies. Point (ii) is both interesting and surprising, suggesting that this aspect of clitic positioning is determined not by the syntax, but by an autonomous morphological component that immediately precedes PF (Halle and Marantz 1993). The possibility will not be pursued here, however.

Another phonological effect within the cluster affects not only the order of clitics, but their shape: the choice between *am* and *mo* for 1/2pl. Although it is not clear from the template in (259), when only one 1/2pl clitic is used in a sentence, *mo* must be used, not *am*:

(308)   a.   *kitab   <u>am</u>

                 book   *1/2PL*

                 'our book; your (pl) book'

      b.   kitab   <u>mo</u>

                 book   *1/2PL*

                 'our book; your (pl) book' (Tegey 1977: 191)

The choice between these forms is also determined by dialect considerations, which will not be discussed here. Tegey (1977: 182–183, 190–192) is best consulted for further information.

There is another interesting phenomenon, apparently occurring at PF, in which the expected order of clitics is disrupted. As was mentioned above in the derivation of (282) from (281), the 3sg clitic *yee* commonly attaches to a preceding word and alters its own shape. Consider the following examples:

(309)   a.   topak-<u>ee</u>   raaworr-e

                 gun(MASC)-3SG   *brought-MASC3SG*

                 'He brought a gun'

b. kheza-y   we-tessteed-a

  woman(ABS)-3SG *PERF-fled-FEM3SG*

  'He/she fled the woman'

c. t-ee   peezan-ee

  PN2SG-3SG *know-2SG*

  'You know him/her' (Tegey 1977: 208)

The presumed underlying forms are as follows:

(310) a. topak   yee raaworr-e

      gun(MASC) *3SG* *brought-MASC3SG*

      'He brought a gun'

    b. kheza   yee we-tessteed-a

      woman(ABS) *3SG* *PERF-fled-FEM3SG*

      'He/she fled the woman'

    c. t   yee peezan-ee[75]

      PN2SG *3SG* *know-2SG*

      'You know him/her'

The phonological alternations of this clitic may be stated thus (formalized from the description of

Tegey 1977: 183):

---

[75] The schwa of the 2sg strong pronoun *te*—not to mention other schwa-final pronouns—is probably epenthetic.

(311)   /yee/   ⇒   [ee]   /   C ___

                      [yee]   V$_{[+stress]}$ ___

                      [y]   elsewhere

The template in (259) indicates that *kho* 'indeed' is the first member of the 2P clitic cluster. Strikingly, however, 3sg *yee* may precede *kho* 'indeed', as long as *yee* incorporates with its host according to the rule in (311). One such paradigm is given below; see Tegey (1977: 209 ff.) for others.

(312)   a.   te   kho   yee   peezan-ee

                PN2SG   *indeed* *3SG*   *know-2SG*

          b.   *te   yee   kho   peezan-ee

                PN2SG   *3SG*   *indeed* *know-2SG*

          c.   t-ee   kho   peezan-ee

                PN2SG-3SG   *indeed* *know-2SG*

                'You indeed know him' (Tegey 1977: 209)

Neither 1sg *mee* nor 2sg *dee*, though similar in form to 3sg *yee*, may undergo this rule, suggesting that the initial consonant of 3sg *yee* is epenthetic, whereas the initial consonants of 1sg *mee* and 2sg *dee* are underlying (Farooq Babrakzai, p.c.).[76]

        Assuming that this process is strictly phonological, it seems that whatever processes give the effect of the template in (259) (argued in this chapter to be the ordinary structures supplied by

---

[76] See Lorimer (1915: 198) for independent evidence that [y] is epenthetic; cf. Babrakzai (1999: 93).

the syntax), some of them must apply only after *yee* contraction (311) has applied (Tegey 1977: 212). In the earlier framework assumed by Tegey, this resulted in the undesirable conclusion that a syntactic process (clitic placement and ordering) must apply after a phonological one—something unexpected in a strictly derivational grammar, with PF rules applying after all others. On the other hand, the analysis offered in this chapter handles this variation without difficulty. The order of clitics in (312b) is ungrammatical at LF, because this order could only have been generated by wrongly merging CliticP (a projection bearing person and number features) above the position in whose specifier the adverb 'indeed' appears. While this same order of clitics is grammatical in (312c), its LF form is presumably as in the canonical (312a); the fact that 3sg *yee* has undergone the PF process in (311) demonstrates that the reordering of 3sg *yee* with respect to *kho* 'indeed' must have occurred at PF, where LF considerations are irrelevant.

Although there is much more that could be said about morphophonological aspects of clitic placement, the conclusion is, once again, that a derivational model of grammar need not be reorganized in order to accommodate the facts of clitic placement in Pashto. Rather, once the independently required principles of grammar are formulated properly, both the positioning of clitics in the clause and most of their ordering with respect to each other are derived without language-specific stipulations—exactly the result desired from a theory of Universal Grammar.

References

Ackerman, Farrell, and Gert Webelhuth. 1998. *A theory of predicates*. Stanford: CSLI Publications.

Aissen, Judith. 1997. On the syntax of obviation. *Language* 73: 705–750.

Alexiadou, Artemis, and Chris Wilder, eds. 1998. *Possessors, predicates, and movement in the determiner phrase*. Amsterdam: John Benjamins Publishing Company.

Babrakzai, Farooq. 1999. Topics in Pashto syntax. Doctoral dissertation, University of Hawai'i at Manoa.

Baker, Mark C. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16: 373–415.

———. 1996. *The polysynthesis parameter*. New York: Oxford University Press.

Becka, Jirí. 1969. *A study in Pashto stress*. Prague: Oriental Institute in Academia.

Bell, Alan, and Mohamad M. Saka. 1983. Reversed sonority in Pashto initial clusters. *Journal of Phonetics* 11: 259–275.

Belletti, Adriana, and Luigi Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language and Linguistic Theory* 6: 291–352.

Bittner, Maria, and Ken Hale. 1996. Ergativity: toward a theory of a heterogeneous class. *Linguistic Inquiry* 27: 531–604.

Bonet, Eulàlia. 1995. Feature structure of Romance clitics. *Natural Language and Linguistic Theory* 13: 607–647.

Boškovic, Zeljko, and Daiko Takahashi. 1998. Scrambling and last resort. *Linguistic Inquiry* 29: 347–366.

Broadwell, George Aaron. 1990. Extending the binding theory: a Muskogean case study. Doctoral dissertation, University of California, Los Angeles.

Bromberger, Sylvain, and Morris Halle. 1989. Why phonology is different. *Linguistic Inquiry* 20: 51–70.

Bubenik, Vit. 1994. On Wackernagel's Law in the history of Persian. *Indogermanische Forschungen* 99: 105–122.

Butt, Miriam. 1995. *The structure of complex predicates in Urdu*. Stanford: CSLI Publications.

Chomsky, Noam. 1970. Remarks on nominalization. In *Readings in English transformational grammar*, ed. Roderick A. Jacobs and Peter S. Rosenbaum, 184–221. Waltham, MA: Ginn and Company.

———. 1981. *Lectures on government and binding: the Pisa lectures*. Dordrecht: Foris Publications.

———. 1995. *The minimalist program*. Cambridge: MIT Press.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. Cambridge: MIT Press, 1991.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective*. New York: Oxford University Press.

Czaykowska-Higgins, Ewa, and M. Dale Kinkade. 1997. Salish languages and linguistics. In *Salish languages and linguistics: theoretical and descriptive perspectives*, ed. Ewa Czaykowska-Higgins and M. Dale Kinkade, 1–68. Berlin: Mouton de Gruyter.

Darmesteter, James. 1888–1890. *Chants populaires des Afghans*. Paris: Imprimerie nationale, E. Lerous.

Davis, Henry. 1995. The clitic-affix distinction in radical head marking languages. In *Proceedings of the 31st Regional Meeting of the Chicago Linguistic Society, Parasession on Clitics*, ed. A. Dainora et al., 64–78. Chicago: Chicago Linguistic Society.

———. 1996. On agreement in St'át'imcets (Lillooet Salish). In *Actes du Deuxième Colloque de Langues et Grammaires*, ed. Lea Nash, Georges Tsoulas, and Anne Zribi-Hertz, 79–94. Paris: Université de Paris VIII.

———. 1998. Subject inflection in Salish. Ms., University of British Columbia.

Demirdache, Hamida. 1997. Dislocation, resumption and weakest crossover. In *Materials on left dislocation*, ed. Elena Anagnostopoulou, Henk van Riemsdijk, and Frans Zwarts, 193–231. Amsterdam: John Benjamins Publishing Company.

Déprez, Viviane. 1998. Semantic effects of agreement: the case of French past participle agreement. *Probus* 10: 1–65.

Dwivedi, Veena. 1994. Syntactic dependencies and relative phrases in Hindi. Doctoral dissertation, University of Massachusetts, Amherst.

Elfenbein, Josef. 1997. Pashto phonology. In *Phonologies of Asia and Africa (including the Caucasus)*, vol. 2, ed. Alan S. Kaye, 733–760. Winona Lake, IN: Eisenbrauns.

Fernández, Beatriz. 1999. On split ergativity: evidence from Basque. *MIT Working Papers in Linguistics* 34: 177–190.

Franco, Jon. 1991. Spanish object clitics as verbal agreement morphemes. *MIT Working Papers in Linguistics* 14: 99–113.

———. 1993. On object agreement in Spanish. Doctoral dissertation, University of Southern California.

Franks, Steven L. 2000. Clitics at the interface. In *Clitic phenomena in European languages*, ed. Frits Beukema and Marcel den Dikken, 1–46. Amsterdam: John Benjamins Publishing Company.

Goodall, Grant. 1987. *Parallel structures in syntax: coordination, causatives, and restructuring*. Cambridge: Cambridge University Press.

Hale, Kenneth, and Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 53–109. Cambridge: MIT Press.

Hale, Ken, and Luciana Storto. n.d. [1996?]. Agreement and spurious antipassives. Ms., MIT.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 111–176. Cambridge: MIT Press.

Halpern, Aaron. 1995. *On the placement and morphology of clitics*. Stanford: CSLI Publications.

Han, Chung-hye. 1998. The structure and interpretation of imperatives: mood and force in Universal Grammar. Doctoral dissertation, University of Pennsylvania.

Hendrick, Randall. 1991. The morphosyntax of aspect. *Lingua* 85: 171–210.

Heston, W. L. 1987. Pashto ambipositions and historical antecedents. In *Selected papers from SALA-7: South Asian Languages Analysis roundtable conference held in Ann Arbor, Michigan, May 17–19, 1985*, ed. Elena Bashir, Madhav M. Deshpande, and Peter Edwin Hook, 163–181. Bloomington: Indiana University Linguistics Club.

Huang, C.-T. James. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531–574.

Iatridou, Sabine, Elena Anagnostopoulou, and Roumyana Izvorski. To appear. An inquiry into the form and meaning of the perfect. In *Ken Hale: a life in language*, ed. Michael Kenstowicz. Cambridge: MIT Press.

Jaeggli, Osvaldo. 1982. *Topics in Romance syntax*. Dordrecht: Foris Publications.

Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2: 39–76.

Kaisse, Ellen M. 1981. Separating phonology from syntax: a reanalysis of Pashto cliticization. *Journal of Linguistics* 17: 197–208.

———. 1985. *Connected speech: the interaction of syntax and phonology*. Orlando: Academic Press.

Karimi, Simin. 1997a. Persian complex verbs: idiomatic or compositional? *Lexicology* 3: 273–318.

———. 1997b. Persian complex predicates and LF incorporation. In *CLS 33: Papers from the Main Session*, ed. Kora Singer, Randall Eggert, and Gregory Anderson, 215–229. Chicago: Chicago Linguistic Society.

Kayne, Richard S. 1989a. Facets of Romance past participle agreement. In *Dialect variation and the theory of grammar*, ed. Paola Benincà, 85–103. Dordrecht: Foris Publications.

———. 1989b. Null subjects and clitic climbing. In *The null subject parameter*, ed. Osvaldo Jaeggli and Kenneth J. Safir, 239–261. Dordrecht: Kluwer Academic Publishers.

Kitahara, Hisatsugu. 1995. Target α: deducing strict cyclicity from derivational economy. *Linguistic Inquiry* 26: 47–77.

Landau, Idan. 1999. Possessor raising and the structure of VP. *Lingua* 107: 1–37.

van der Leeuw, Frank. 1995. Alignment and integrity constraints in cliticization. In *Proceedings of the 31st Regional Meeting of the Chicago Linguistic Society, Parasession on Clitics*, vol. 2, ed. A. Dainora et al., 168–180. Chicago: Chicago Linguistic Society.

———. 1997. *Clitics: prosodic studies*. The Hague: Holland Academic Graphics. [Doctoral dissertation, University of Amsterdam.]

Lorimer, D. L. R. 1915. *Pashtu part 1: syntax of colloquial Pashtu, with chapters on the Persian and Indian elements in the modern language*. Oxford: Clarendon Press.

MacKenzie, D. N. 1987. Pashto. In *The world's major languages*, ed. Bernard Comrie, 547–565. New York: Oxford University Press.

Mahajan, Anoop. 1990. The A/A-bar distinction and movement theory. Doctoral dissertation, MIT.

McCloskey, James, and Kenneth Hale. 1983. On the syntax of person-number inflection in Modern Irish. *Natural Language and Linguistic Theory* 1: 487–533.

Miller, George A., and Noam Chomsky. 1963. Finitary models of language users. In *Handbook or Mathematical Psychology*, vol. 2, ed. R. Duncan Luce, Robert R. Bush, and Eugene Galanter, 419–491. New York: John Wiley and Sons.

Miyagawa, Shigeru. 1997. Against optional scrambling. *Linguistic Inquiry* 28: 1–25.

———. To appear. EPP, scrambling and *wh*-in-situ. In *Ken Hale: a life in language*, ed. Michael Kenstowicz. Cambridge: MIT Press.

Mohammad, Jan. 1993. Second-position clitics in Pashto. Ms., University of Arizona.

Morgenstierne, Georg. 1927. An etymological vocabulary of Pashto. *Skrifter Det Norske Videnskaps-Akademi i Oslo*, 3.

Munro, Pamela. 1984. The syntactic status of object possessor raising in Western Muskogean. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, ed. Claudia Brugman and Monica Macaulay, 634–649. Berkeley: Berkeley Linguistics Society, University of California.

Murasugi, Kumiko G. 1997. Is there an ergative parameter? In *CLS 33: Papers from the Panels*, ed. Kora Singer, Randall Eggert, and Gregory Anderson, 239–251. Chicago: Chicago Linguistic Society.

Myers, Scott. 1997. OCP effects in optimality theory. *Natural Language and Linguistic Theory* 15: 847–892.

Newmeyer, Frederick J. 1986. *Linguistic theory in America*. 2nd ed. San Diego: Academic Press.

Payne, Doris L. and Immanuel Barshi, eds. 1999a. *External possession*. Amsterdam: John Benjamins Publishing Company.

———. 1999b. External possession: what, where, how, and why. In *External possession*, ed. Doris L. Payne and Immanuel Barshi, 3–29. Amsterdam: John Benjamins Publishing Company.

Penzl, Herbert. 1955. *A grammar of Pashto: a descriptive study of the dialect of Kandahar, Afghanistan*. Washington, D.C.: American Council of Learned Societies.

Perlmutter, David M. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Winston.

Pesetsky, David. 1997. Optimality theory and syntax: movement and pronunciation. In *Optimality theory: an overview*, ed. Diana Archangeli and D. Terence Langendoen, 134–170. Malden, MA: Blackwell Publishers.

———. 1998. Some optimality principles of sentence pronunciation. In *Is the best good enough?: optimality and competition in syntax*, ed. Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky, 337–383. Cambridge: MIT Press.

Progovac, Ljiljana. 1993. Clitics in Serbian/Croatian: Comp as the second position. Ms., Wayne State University.

———. 1999. Clitic second and verb second. Ms., Wayne State University.

Ramchand, Gillian Catriona. 1991. Complex predicate formation in Bangla. In *The proceedings of the ninth West Coast Conference on Formal Linguistics*, ed. Aaron L. Halpern, 443–458. Stanford: CSLI Publications.

———. 1997. *Aspect and predication: the semantics of argument structure*. Oxford: Clarendon Press.

———. 1998. Deconstructing the lexicon. In *The projection of arguments: lexical and compositional factors*, ed. Miriam Butt and Wilhelm Geuder, 65–96. Stanford: CSLI Publications.

Riemsdijk, Henk van. 1997. Left dislocation. In *Materials on left dislocation*, ed. Elena Anagnostopoulou, Henk van Riemsdijk, and Frans Zwarts, 1–10. Amsterdam: John Benjamins Publishing Company.

Roberts, Taylor. 1997. The optimal second position in Pashto. In *Phonology in Progress – Progress in Phonology: HIL Phonology Papers III*, ed. Geert Booij and Jeroen van de Weijer, 367–401. The Hague: Holland Academic Graphics.

Romano, Christine. 1991. Mixed headedness in American Sign Language: evidence from functional categories. *MIT Working Papers in Linguistics* 14: 241–254.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT. [Pp. references are to the edition published as *Infinite syntax!*, Norwood, NJ: Ablex Publishing, 1986.]

———. 1972. Doubl-ing. *Linguistic Inquiry* 3: 61–86.

Rudin, Catherine. 1997. AgrO and Bulgarian pronominal clitics. In *Annual workshop on formal approaches to Slavic linguistics: the Indiana meeting, 1996*, ed. Martina Lindseth and Steven Franks, 224–252. Ann Arbor: Michigan Slavic Publications.

Sapir, Edward. 1930. *Southern Paiute, a Shoshonean language*. In Sapir (1992).

———. 1992. *Southern Paiute and Ute: linguistics and ethnography. The collected works of Edward Sapir*, vol. 10, ed. William Bright. Berlin: Mouton de Gruyter.

Schütze, Carson. 1994. Serbo-Croatian second position clitic placement and the phonology-syntax interface. *MIT Working Papers in Linguistics* 21: 373–473.

Selkirk, Elisabeth. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge: MIT Press.

———. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3: 371–405.

Sells, Peter. 1984. Syntax and semantics of resumptive pronouns. Doctoral dissertation, University of Massachusetts, Amherst.

Shafeev, D. A. 1964. *A short grammatical outline of Pashto*. Ed. and trans. Herbert H. Paper. Bloomington: Indiana University; and The Hague: Mouton and Co.

Shlonsky, Ur. 1992. Resumptive pronouns as a last resort. *Linguistic Inquiry* 23: 443–468.

Skjærvø, Prods O. 1989. Pashto. In *Compendium linguarum Iranicarum*, ed. Rudiger Schmitt, 384–410. Wiesbaden: Reichert.

Sportiche, Dominique. 1996. Clitic constructions. In *Phrase structure and the lexicon*, ed. Johan Rooryck and Laurie Zaring, 213–276. Dordrecht: Kluwer Academic Publishers.

Stjepanovic, Sandra. 1998. On the placement of Serbo-Croatian clitics: evidence from VP-ellipsis. *Linguistic Inquiry* 29: 527–537.

Suñer, Margarita. 1988. The role of agreement in clitic-doubled constructions. *Natural Language and Linguistic Theory* 6: 391–434.

Tegey, Habibullah. 1977. The grammar of clitics: evidence from Pashto and other languages. Doctoral dissertation, University of Illinois at Urbana-Champaign.

———. 1979. Ergativity in Pushto (Afghani). In *Linguistic method: essays in honor of Herbert Penzl*, ed. Irmengard Rauch and Gerald F. Carr, 369–418. The Hague: Mouton Publishers.

Tegey, Habibullah, and Barbara Robson. 1996. *Pashto reference grammar*. Washington, D.C.: Center for Applied Linguistics.

Tomic, Olga Mišeska. 1996. The Balkan Slavic clausal clitics. *Natural Language and Linguistic Theory* 14: 811–872.

Travis, Lisa deMena. 1984. Parameters and effects of word order variation. Doctoral dissertation, MIT.

Ura, Hiroyuki. 1996. Multiple feature-checking: a theory of grammatical function splitting. Doctoral dissertation, MIT.

Van Gelderen, Elly. 1997. Universals and minimalist features: checking in AgrO. In *Studies on universal grammar and typological variation*, ed. Artemis Alexiadou and T. Alan Hall, 181–196. Amsterdam: John Benjamins Publishing Company.

VanLoon, Jennifer. 1997. Pronominal morphology in Sulimani Kurdish. In *Yale a-morphous linguistics essays: studies in the morphosyntax of clitics*, ed. Lizanne Kaiser, 151–174. New Haven: Dept. of Linguistics, Yale University.

Williams, Edwin. 1977. Across-the-board application of rules. *Linguistic Inquiry* 8: 419–423.

———. 1978. Across-the-board rule application. *Linguistic Inquiry* 9: 31–43.