

Article

"Berber Clitic Doubling and Syntactic Extraction"

Mohamed Guerssel

Revue québécoise de linguistique, vol. 24, n° 1, 1995, p. 111-133.

Pour citer cet article, utiliser l'information suivante :

URI: <http://id.erudit.org/iderudit/603105ar>

DOI: 10.7202/603105ar

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : info@erudit.org

BERBER CLITIC DOUBLING AND SYNTACTIC EXTRACTION*

Mohamed Guerssel
Université du Québec à Montréal

1. Introduction

THE SUBJECT markers of Berber have traditionally been considered to be agreement markers. I will argue in this paper that they are actually clitics, and that subject-verb agreement has no expression other than that of the clitic.¹ Furthermore, I will argue that the expression of overt lexical subjects in Berber is an instance of clitic doubling. As is well known, clitic doubling constructions resist syntactic extraction. Various explanations for the alleged impossibility of extraction have in the past been put forth. Based on evidence from Berber, I will argue that syntactic extraction *is* possible out of doubling constructions, provided a default clitic is available in the language. More precisely it will be shown that when extraction takes place out of the subject position, whether it is relativization, clefting, or *wh*-movement, a clitic is still present, but it has a default value for person, number, and gender features. The

* In writing this paper I have benefited from the comments of the following: Isabelle Haik, Ken Hale, Louise Lavoie, Beth Levin, Victor Manfredi, Jamal Ouhalla, Yves Roberge, Ur Shlonsky, Lisa Travis, Laurie Tuller and an anonymous reviewer for this journal. Preparation of this manuscript was supported by SSHRC grant #410-91-0716.

¹ The main dialect described in this paper is that of Ait Seghrouchen, spoken in the Middle Atlas in Morocco. The following notation is used: [e] is an excrescent vowel in Ait Seghrouchen, and a schwa in the other dialects; c = voiceless palatal fricative; j = voiced palatal fricative; gh = voiced uvular fricative; ttc = geminate voiceless palatal affricate. A glide is pronounced as its corresponding vowel when nonadjacent to a vowel. All other symbols are standard. The following abbreviations are used: The numbers 1, 2, and 3 stand for first, second, and third person, respectively; s = singular, ms = masculine singular, mp = masculine plural, fs = feminine singular, fp = feminine plural; obj = object, dat = dative, inf = infinitive.

default clitic is utilized regardless of the person, number, gender features of the DP it is associated with. The agreement peculiarity will be exploited for the purpose of providing a generalized account of the clitic doubling phenomenon. More precisely, the notion of clitic chain will be introduced and defined as the conjunction of a clitic and the DP it is coindexed with. A clitic chain is said to be well-formed provided its two members agree in terms of ϕ -features. If a trace is assumed to have a default value for ϕ -features, then extraction out of a doubling construction is possible provided a default clitic forms a clitic chain with the trace at S-structure. A comparison of those Berber dialects which exhibit clitic doubling and those that do not reveals that the two types of dialects contrast in other aspects of the grammar, exactly as they are expected to.

2. On a Subject-Object Asymmetry

Berber is a VSO Null Subject language.² The subject may or may not be represented by an overt DP. Information about the missing subject is presumably expressed by one or more verbal affixes. These affixes convey inflectional information about person, number and gender. For example, each string in (1a) is a complete, well-formed sentence. In (1b) are sentences where the subject is realized, and in (1c) appear sentences where both a subject and a direct object are realized.

- | | | |
|--------|--|---------------------------|
| (1) a. | y-swu.
3ms-drank | ‘He drank.’ |
| | t-swu.
3fs-drank | ‘She drank.’ |
| b. | y-swu wryaz.
3ms-drank man | ‘The man drank.’ |
| | t-swu tfunast.
3ms-drank cow | ‘The cow drank.’ |
| c. | y-wtu wryaz tafunast.
3ms-hit man cow | ‘The man hit the cow.’ |
| | t-wtu tfunast aryaz.
3fs-hit cow man | ‘The cow kicked the man.’ |

²The sentences in (1c) are representative of the surface VSO order of Berber. The question of word order will not be addressed in this paper since it has no bearing on the issues to be raised. I will simply assume that the basic order is SVO, and indicate it as such whenever a D-structure is given.

Berber appears to be similar to languages like Italian, Russian, and Spanish, which permit the subject position to be empty, and different from a language like English, where the subject position must be realized. What distinguishes the two types of languages is that the subject position in Null Subject languages may be occupied by the empty category *pro*, cf. Chomsky (1982), Rizzi (1986a, 1986b) among others. In languages like English, the subject position may not be phonetically empty.

Setting aside the Null Subject issue, we examine a particular property of overt Berber subject DPs. What is remarkable about overt Berber subjects is that they may only have a definite interpretation, in contrast with object DPs which may be either definite or indefinite.³ Thus, although the translations given in (1) above are correct, they are not exactly accurate. A more accurate rendition of the sentences in (1b-c) is given in (2).

- (2) a. *y-swu wryaz.* 'The man drank.' / '*A man drank.'
 3ms-drank man
- t-swu tfunast.* 'The cow drank.' / '*A cow drank.'
 3ms-drank cow
- b. *y-wtu wryaz tafunast.* 'The man hit the cow.' / 'The man hit a cow.'
 3ms-hit man cow '*A man hit a cow.' / '*A man hit the cow.'
- t-wtu tfunast aryaz.* 'The cow kicked the man.' / 'The cow kicked a man.'
 3fs-hit cow man '*A cow kicked a man.' / '*A cow kicked the man.'

Although the object is ambiguous as to whether it is definite or indefinite, the subject may only be definite. This is indeed a very unusual feature. Given the set of forms represented in (2) an important question arises: How is an indefinite overt subject expressed in Berber? In many languages that do not exhibit overt articles, indefiniteness is typically expressed by means of the numeral 'one'. This is true of Berber. The numerals *ijj* 'one' (masculine) and *ict* 'one' (feminine) are utilized to express indefiniteness or specificity. But this fact sheds no light on the obligatory interpretation of the subjects in (2) as definite, simply because the numeral 'one' may also be used with objects, yielding an indefinite or specific interpretation. The question raised with respect to the contrast in (2) remains posed. Why is the interpretation of the object ambiguous with respect to definiteness, whereas it is uniquely definite with respect to subjects?

³ What is even more remarkable is that no one, to the best of my knowledge, has ever brought out this particular feature of Berber, and put forward an explanation for it.

Close examination of the two DP types reveals that subjects differ from objects in terms of their morphological shape. What bearing, if any, does this contrast have on the definiteness issue? As we show in what follows, there is no direct connection between the morphological shape of a DP and definiteness. Before doing so, we must identify the morphological contrast in question.

As has long been recognized, the morphological shape of a noun (or a DP) in Berber is determined by the syntactic context in which it appears. The alternation is between what has come to be known as the Construct State and the Free State. Post-verbal subjects and the object of the majority of prepositions are in the Construct State. Direct objects, left-dislocated DPs, as well as the object of a limited number of prepositions are in the Free State. This contrast is highlighted in (3).⁴

(3)		<i>Construct</i>	<i>Free</i>	
	<i>Masculine</i>	wryaz	aryaz	'man'
	<i>Feminine</i>	tcurt	tacurt	'ball'

Are we to conclude that the Construct State is to be correlated with definiteness, while the Free State goes hand in hand with an indeterminate value for definiteness? The answer to this question is clearly negative, in view of the fact that the object of a preposition is also in Construct, but behaves like a direct object in terms of definiteness. That is, it is ambiguous between a definite and an indefinite reading, as illustrated in (4).

- (4) a. y-wcu wryaz tacurt i wrba. 'The man gave a ball/the ball to a boy/the boy.'
3ms-gave man ball to boy
- b. y-ssers wrba azru x tcurt. 'The boy put a stone/the stone on a ball/the ball.'
3ms-put boy stone on ball
- c. y-ssers wrba tacurt x wzru. 'The boy put a ball/the ball on a stone/the stone.'
3ms-put boy ball on stone
- d. t-wtu tcurt arba. 'The ball hit a boy/the boy.'
3fs-hit ball boy
- e. y-wda wzru. 'The stone fell.'
3ms-fell stone

The Construct State of the nominals in (4) are *wryaz* 'man', *wrba* 'boy', *tcurt* 'ball', and *wzru* 'stone'. Their respective Free State forms are *aryaz*, *arba*, *tacurt*, and *azru*. What these new facts clearly show is that there is no direct

⁴For the phonological aspects of the Construct phenomenon, see Guerssel (1983).

correlation between State and definiteness, contrary to what the data in (2) might lead us to believe. Although the subject is in Construct, and uniquely definite, the direct and indirect objects are both indeterminate with respect to indefiniteness in spite of the fact that one of them is in Construct but the other one must be in the Free State. Examination of the syntactic distribution of the various DPs, as illustrated in (2) and (4) reveals that there is no connection between the Construct State and definiteness. As a consequence, the question raised earlier remains posed: Why is it that overt subject DPs in Berber may only be definite? In order to answer this question, it is perhaps worth examining other structures involving DPs, where such DPs are required to be definite. An attempt to achieve this goal is made in the following section.

3. Clitic Doubling

Recall from the examples given in (4) that dative DPs in Berber may be interpreted either as definite or indefinite. There is, however, a construction where the dative DP must be definite. This fact is illustrated in (5).⁵

- (5) a. Wci-x aysum i-wmucc.
gave-1s meat to-cat
'I gave meat to the cat. / I gave meat to a cat.'
- b. Wci-x-as aysum.
gave-1s-3s:dat meat
'I gave him meat.'
- c. Wci-x-as_i aysum i-wmucc_i.
gave-1s-3s:dat meat to-cat
'I gave meat to the cat.' [Literally, 'I gave him_i meat to the cat_i.']
* 'I gave meat to a cat.'

The indirect object in (5a) is a full DP, marked by *i*, and is understood either as a definite or an indefinite DP. In (5b) there is a dative clitic but no overt dative DP. In (5c) the interpretation of the dative DP must be definite. As is well known, definiteness is a hallmark of clitic doubling constructions. Examination of the structure in (5c) reveals that we are indeed in the presence of clitic doubling. Both the dative clitic and the corresponding full DP are present. Two

⁵In (5) the direct object may be definite or indefinite, as already shown in the text. This feature is ignored in the translations since it is not relevant for the discussion. Objects are arbitrarily given as indefinite.

simple examples from languages that allow clitic doubling provide evidence that definiteness goes hand in hand with clitic doubling. In River Plate Spanish, objects may be clitic doubled and the object must be definite. In Colloquial French, subject clitic doubling is possible, but again the subject must be definite. The first three French sentences are taken from Roberge (1986).

(6) a. River Plate Spanish

Ví el chico. 'I saw the boy.'

I-saw the boy

Ví un chico. 'I saw a boy.'

I-saw the boy

Lo ví a el chico. 'I saw the boy.' [Literally, 'I saw him to the boy']
him I-saw to a boy.

*Lo ví a un chico. [Literally, 'I saw him to a boy.']
him I-saw to a boy.

b. Colloquial French (Roberge 1986, p. 166)

Son ami est toujours là. 'His friend is always there.'

his friend is always there

Un ami est toujours là. 'A friend is always there.'

a friend is always there

*Un ami il est toujours là. [Literally, 'A friend he is always there.']
a friend he is always there

Son ami il est toujours là. [Literally, 'His friend he is always there.']
his friend he is always there

The doubled Berber DP in (5c) is analogous to the doubled DPs in (6). In each set, both a clitic and a lexical DP are present.

Following Roberge (1986), I will adopt the view that the presence or absence of a lexical DP in doubling constructions is dictated by Case theory. More specifically, I will assume that clitics may or may not absorb Case. If the clitic absorbs Case, then the canonical DP-position coindexed with that clitic must be *pro*. If it does not, then an overt lexical DP will have to be generated in that position. The structures of the Berber VPs in (5a-c) are given in (7a-c).

- (7) a. [_{VP} Verb DP i-DP]
 b. [_{VP} Verb-clitic_i DP *pro*_i]
 c. [_{VP} Verb-clitic_i DP i-DP_i]

If, as assumed, clitics may or may not absorb Case, then the above three structures will be derived in a straightforward manner. (7a) does not contain a clitic and poses no problem. If a clitic is present and the option of absorbing Case is taken, then the dative position must be *pro*, as in (7b), otherwise a Case Filter violation ensues. If the clitic does not absorb Case, then the dative position must contain a lexical DP, as in (7c). Crucially, however, the clitic-doubled DP must be definite, as in (5c). In the dative construction, Case is presumably assigned to the doubled lexical DP by the marker [i]. Similarly, a preposition heads a DP object in the Spanish and French examples in (6), just in case the construction involves doubling.

Analyzing the overt realization of Berber lexical subjects as an instance of clitic doubling provides an explanation for the question raised in section 1 relating to the obligatory interpretation of Berber subjects as definite. If Berber lexical subjects are treated as instances of clitic doubling, then they fall under the general phenomenology of clitic doubling whereby clitic-doubled DPs are necessarily definite.⁶

The asymmetry between subjects and objects in Berber regarding definiteness is no longer a mystery. Overt subjects are instances of clitic doubling, overt objects are not. We are now in a position to examine another asymmetry between subjects and objects, the one pointed out in section 1 regarding the State phenomenon, cf. the data in (2), (3) and (4).

The analysis provided in Guerssel (1992) of the State phenomenon departs radically from previous approaches. There it is demonstrated that with the exception of two items, all the lexical items traditionally considered to be prepositions in Berber are in fact case markers. Together with their DP complements they form Kase Phrases. What has traditionally been treated as the Construct State is either a Determiner Phrase or a Kase Phrase where the head K is not realized. There are thus only three shapes that the maximal projection of a KP may take. These are illustrated in (8) with the nominal *azru* 'stone'.⁷

- (8) a. [_{KP} s [wzru]] 'with a stone' (Case marker: instrumental *s*)
 b. [_{KP} a [zru]] 'stone' (Case marker: default *a*)
 c. [_{KP} ∅ [wzru]] 'stone' (Case marker: empty)

⁶ As argued by May (1985), indefinite NPs, unlike definite NPs, undergo Quantifier Raising. As we will see later, clitic doubling resists extraction. Presumably, then, indefinite NPs are disallowed in doubling constructions because that would imply extraction at LF.

⁷ For the absence of the determiner *w* in (8b), see Guerssel (1992).

There are close to a dozen case markers, including instrumental (illustrated in (8a)), genitive, comitative, superessive, elative, and so on. With a DP as a complement, each case marker heads a Kase Phrase of the type shown in (8a).

Two markers that are absent from the inventory of Berber case markers are the accusative and oblique markers. As a consequence, what is termed a default marker, illustrated in (8b), marks the object of a verb as well as the object of a true preposition. This position is defended in Guerssel (1992, p. 187-191). The default case marker also marks dislocated KPs, be they subjects or objects. This is in line with standard assumptions whereby abstract Case is assigned by default to KPs generated outside IP.⁸

The last KP type represented in (8c), where the K position is empty, is syntactically instantiated in only one position: the post-verbal position. Subject KPs thus contrast with all the other KP types where Case is morphologically realized as K. On the basis of the analysis proposed in the foregoing, subject KPs contrast with other KP types in yet another way. They are instances of clitic doubling. There is an explanation available for the absence of a case mark on subjects if one simple assumption is made: Nominative Case is realized in AGR as a clitic, and hence may not also be realized on the KP subject. Case marking is thus a property of the clitic chain as a whole. This assumption is compatible with the observation that left-dislocated subjects appear with the default marker, as exemplified in (8b), and are therefore morphologically indistinguishable from objects. This is to be expected since they do not form a clitic chain with the clitic. In left dislocation the clitic chain consists of a clitic and the empty category *pro*.

The proposed analysis makes a strong prediction concerning the subject-object asymmetry as it relates to case marking. The prediction is that in a clitic doubling construction involving an object, the object would exhibit the pattern in (8c), just as clitic doubled subjects do, instead of the pattern (8b). As will be shown in what follows, this prediction is borne out by the facts of a dialect that displays clitic doubling of objects.

Although direct object doubling is not instantiated in the Ait Seghrouchen dialect of Berber discussed so far, there are Berber dialects where it is attested. One such dialect is Taqbaylit, spoken in the Kabylie region of Algeria. Taqbaylit is similar to the Ait Seghrouchen dialect in all the syntactic aspects discussed

⁸In Classical Arabic, for instance, left dislocation is achieved by means of one of two strategies: movement or base-generation outside IP. In movement constructions the moved DP exhibits the case mark corresponding to the case value it is normally assigned in its canonical position. In the other case, it is uniquely marked with the nominative case marker. Nominative is the default value in Classical Arabic.

so far. In addition, it allows clitic doubling of direct objects. A typical paradigm appears in (9).

- (9) a. wala-gh amci.
 saw-1s cat
 'I saw the cat. / 'I saw a cat.'
- b. wala-gh-t.
 saw-1s-3ms-obj
 'I saw it.'
- c. wala-gh-t_i wemci_i.
 saw-1s-3ms:obj cat
 'I saw the cat.' / '*I saw a cat.' [Literally, 'I saw it_i the cat_i']
- d. *wala-gh-t_i amci_i.
 saw-1s-3ms:obj cat

Again we see that the same definiteness restriction noted earlier holds for clitic doubled objects in Taqbaylit. In (9a) there is no object clitic; the object may either be definite or indefinite. In (9b) there is a clitic but no overt lexical subject. In the clitic doubling construction in (9c), taken from Galand (1979), an object clitic cooccurs with an overt lexical object. The above paradigm and the indirect object paradigm in (5) are clearly parallel. The above paradigm is also parallel to the subject paradigm, except for the fact that subject clitic doubling is obligatory. The behavior of object DPs in Taqbaylit fits in neatly with the rest of Berber DPs.

Examination of the data in (9) reveals a striking contrast between the object DP in (9a) and the one in (9c). The former exhibits default case marking, the latter is unmarked for case. The string in (9c) patterns in exactly the same manner as a clitic doubled subject, in that it is unmarked for case, exactly as expected. In (9a) the object displays default case, again as expected. In (9c), where clitic doubling is involved, the lexical object may not exhibit case since case is realized on the clitic, exactly as was shown to be the case with clitic doubled subjects. For this reason, the string in (9d) is ungrammatical. It involves clitic doubling as well as the realization of default case on the object, two features that are incompatible.

The overall picture that emerges is that overt lexical subjects are manifestations of clitic doubling in Berber. Overt subjects have been shown to exhibit the same properties as those direct objects and indirect objects that are uncontroversially cases of syntactic doubling. May the observed parallelism be extended to other aspects of clitic doubling as, for instance, syntactic

extraction, which is typically blocked when doubling is involved? This question is addressed in the following section.

4. Syntactic Extraction out of Doubling Constructions

As has been observed in the past, syntactic extraction out of clitic doubled DPs is prohibited. As reported in Roberge (1986) for instance, *wh*-extraction of direct objects out of doubling constructions is not permitted in River Plate Spanish, Pied Noir French, and Rumanian. Similarly, Colloquial French, Pied Noir French, Trentino Italian, Fiorentino Italian, all of which exhibit subject clitic doubling, do not allow the *wh*-extraction of clitic-doubled subjects. Berber does behave like the above languages when indirect objects are concerned. Clitic doubled indirect objects may not undergo *wh*-movement, but the nondoubled ones may, as illustrated in (10).

- (10) a. Wci-x aysum i wqqzin.
gave-1s meat to dog
'I gave meat to the dog.'
- b. Mumi_i wci-x aysum t_i?
whom gave-1s meat
'Who did I give meat to?'
- c. Wci-x as_i aysum i wqqzin_i.
gave-1s him:dat meat to dog
'I gave [him] meat to the dog.'
- d. *Mumi_i as_i wci-x aysum t_i?
whom him:dat gave-1s meat [Literally, 'Whom did I give him meat?']

In Taqbaylit Berber, which allows clitic doubling of direct objects, the same restriction holds, as illustrated in (11).

- (11) a. Wala-gh amci.
saw-1s cat
'I saw the cat.'
- b. Acu_i ay wala-gh t_i?
what that saw-1s
'What did I see?'

- c. Wala-gh t_i wemcic_i?
 saw-1s him cat
 'I saw [him] the cat.'
- d. *Acu_i ay t_i wala-gh t_i ?
 what that him saw-1s [Literally, 'What did I see him?']

Several proposals dealing with the impossibility of extraction out of doubling constructions have been made in the past. Some examples appear in Jaeggli (1981), Borer (1983), Aoun (1981), and Roberge (1986).⁹ In spite of the fact that Berber subjects are instances of clitic doubling as we have assumed, they do allow syntactic extraction. In this respect the possibility of *wh*-extraction of Berber subjects contrasts drastically with the impossibility of extraction out of doubling constructions in those languages that allow clitic doubling. Does this asymmetry force us to abandon the claim that Berber overt lexical subjects are instances of clitic doubling? In what follows, we will demonstrate that it does not.

Perhaps the most crucial aspect of the behavior of subject clitics in Berber involves agreement. In Berber declaratives, there is a distinction of person, number, and gender agreement for subjects, as illustrated in the full paradigm exhibited in (12), where the agreement markers are italicized.

(12) <i>ssnw-x</i> .	'I cooked.'
<i>t-ssnw-t</i> .	'You cooked.'
<i>y-ssnw wryaz</i> .	'The man cooked.'
<i>t-ssnw temttutt</i> .	'The woman cooked.'
<i>n-ssnw</i> .	'We cooked.'
<i>t-ssnw-m</i> .	'You (mp) cooked.'
<i>t-ssnw-nt</i> .	'You (fp) cooked.'
<i>ssnw-n iryazn</i> .	'The men cooked.'
<i>ssnw-nt tsednan</i> .	'The women cooked.'

The agreement observed in (12) does not carry over to extraction constructions. In extraction constructions the verb shows no agreement. This lack of agreement is even more striking in relative clauses and clefts than it is in questions, as revealed by a comparison of the following examples with the strings shown in (12).

⁹Roberge (1986) is a particularly excellent study of typological differences regarding doubling.

(13) *Questions:*

w ay y-ssnw-n? ‘Who cooked?’
 who that cooked

Relatives:

aryaz din y-ssnw-n ‘the man who cooked’
 man who cooked
 tarbatt din y-ssnw-n ‘the girl who cooked’
 girl who cooked
 iryazn din y-ssnw-n ‘the men who cooked’
 men who cooked
 tisdnan din y-ssnw-n ‘the women who cooked’
 women who cooked

Clefts:

aryaz ay y-ssnw-n. ‘(It is) the man that cooked.’
 man that cooked
 tarbatt ay y-ssnw-n. ‘(It is) the girl that cooked.’
 girl that cooked
 iryazn ay y-ssnw-n. ‘(It is) the men that cooked.’
 men that cooked
 tisdnan ay y-ssnw-n. ‘(It is) the women that cooked.’
 women that cooked

The canonical shape *y-stem-n* of the verb has traditionally been referred to as the neutral form, or participial form. In the declaratives laid out in (12) each of the verbs bears distinct markers (clitics within the proposed analysis) that express different values for person, number, and gender. The verbs in (13), on the other hand, are not so marked.

The agreement markers used in (12) are incompatible with the question, the relatives and the clefts given in (13). Their use in (13) would lead to ungrammaticality. Likewise, the use of the neutral affixes of (13) in the contexts of (12) would also lead to ungrammaticality. In the verbal forms in (13) agreement is unspecified, so to speak. Now since subject clitics are bearers of agreement, and since there is no agreement in extraction out of subject constructions, I will assume that the discontinuous morpheme *y-n* is a “neutral” clitic, neutral in the sense that it has a default value for ϕ -features. A new factor thus seems to be at play in the dichotomy regarding extraction in clitic doubled constructions: agreement. In what follows I will argue that a proper characterization of the extraction phenomenon is rooted in a proper characterization of agreement.

Before doing so, a particular feature of default agreement will be examined. Consider the two simple English examples given in (14).

- (14) a. Who likes beans?
 b. *Who like beans?

The question in (14a) exhibits third person singular agreement on the verb, and it is well-formed. The sentence in (14b), on the other hand, which does not exhibit such marking, is ungrammatical. Why should this be the case? The answer to this question is straightforward. The marking on the verb of the well-formed structure is not an expression of third person singular agreement, but rather an expression of default agreement. It is generally agreed that in language after language, default agreement marking is typically expressed by a third person agreement marker. We will simply record this feature for now and return to it when appropriate.

Going back to clitics, we submit that the relation holding between a clitic and the DP it is coindexed with is crucially one of agreement. This observation is recorded in (15) where a definition of a clitic chain is given.

(15) *Definition of a clitic chain:* A clitic agrees in person, number, and gender features with the DP it is coindexed with. The clitic and the DP are said to form a clitic chain.

If there is no agreement between the clitic and the DP it is coindexed with, ill-formedness results, as shown in the examples given in (16), from Pied Noir French, a language that allows both subject and object clitic doubling.

- (16) a. Jean_i il_i est parti. 'Jean has left.'
 Marie_i elle_i est partie. 'Marie has left.'
 Les enfants_i ils_i sont partis. 'The children have left.'
 *Paul_i elle_j est partie. [Literally, 'Paul she has left.']
 *Marie_i il_j est parti. [Literally, 'Marie he has left.']
 *L'enfant_i ils_j sont partis. [Literally, 'The child they have left.']
- b. Il l_i'a frappée à la fille_i. 'He hit the girl.'
 Il les_i a frappées aux filles_i. 'He hit the girls.'
 *Il les_i a frappés à la fille_j. [Literally, 'He hit them the girl.']
 *Il l_i'a frappé aux filles_j. [Literally, 'He hit him the girls.']

It may at first seem trivial, but the starred strings shown in (16) are ungrammatical simply because they stand in violation of the definition given in (15). In each case the clitic does not agree with the DP it is coindexed with. This simple observation will be instrumental in ruling out those cases where

ungrammaticality results whenever extraction takes place. More precisely, it will be demonstrated that the ungrammaticality of those strings that are purportedly the result of extraction out of clitic doubling constructions are ungrammatical for the same reason that those shown in (16) are. The two structure types involve ill-formed clitic chains, and that is what leads to ungrammaticality.

Since Berber seems to be singular among clitic doubling languages with respect to subject extraction, it will be dealt with first. Consider the following D-structure.

(17) [_{CP} [_{IP} [_{DP} *wh*] [_{I'} [_I [_{AGR} clitic agr] [tense]]] [VP]]]

Wh-movement yields the following structure.

(18) [_{CP} *wh_i* [_{IP} [_{DP} *t_i*] [_{I'} [_I [_{AGR} clitic_i agr] [tense]]] [VP]]]

The grammaticality of a sentence corresponding to (18) depends on the nature of the subject clitic. If the third person singular clitic is utilized, on the assumption that third singular agreement is the default option, and is normally associated with subject extraction, as was shown to be the case in (14), we would expect grammaticality to result. But as shown in (19a), the derived string is illformed. If on the other hand, the default clitic *y-n* is used, grammaticality is obtained, as illustrated in (19b).¹⁰

(19) a. **w ay y-ssnw?* (phonetically, [wiggessnw])
 who that 3ms-cooked

b. *w ay y-ssnw-n?* (phonetically, [wiggessnwn])
 who that cooked:neutral
 'Who cooked?'

What (19) clearly shows is that (19a) and (14a) are not analogous, in spite of the fact that agreement in both cases is third masculine singular. The difference is obvious. In the English example in (14a) the agreement on the verb is default agreement. In the Berber example in (19a), agreement is expressed by the clitic *y-*. In this case, the clitic is the expression of the third person masculine singular, not default agreement. Default agreement in Berber is instantiated in (19b), and grammaticality results. The grammaticality of (19b) and the ill-formedness of (19a) can be accounted for on the basis of a simple assumption: a DP-trace has a default value for agreement features. If this assumption is adopted, then

¹⁰ That the clitic complex *y-n* has a default value for person, number, and gender, is clearly revealed by the examples given in (13).

the contrast in (19) receives a straightforward explanation. (19a) is ungrammatical because there is a mismatch between the feature values of the members of the clitic chain. (19b) is ruled in because there is a match. The sentence in (19a) is thus ungrammatical for the same reason that the following set consists of ungrammatical sentences.

- (20) *w ay e_i ssnw- x_j ?
 who that cooked-1s [Literally, 'Who I cooked?']
- *w ay e_i t-ssnw-t?
 who that cooked:2s [Literally, 'Who you cooked?']
- *w ay e_i t-ssnw?
 who that 3fs-cooked [Literally, 'Who she cooked?']
- *w ay e_i n-ssnw?
 who that 1mp-cooked [Literally, 'Who we cooked?']
 etc.

In each case in (20) there is a feature mismatch between the members of a clitic chain. The four sentences exhibit respectively the first person singular, second person singular, third person feminine singular, and the first person plural clitics respectively. In each case the other member of the clitic chain, namely the trace, is unmarked for the relevant features. The sentences thus stand in contradiction with the definition of a well-formed clitic chain as defined in (15), and are therefore ruled out. They are ungrammatical for exactly the same reason that the Pied Noir French sentences exhibited in (16) are, independently of extraction. Extraction out of subject doubling is indeed possible, as evidenced by sentence (19b). How about Pied Noir French, where extraction is apparently disallowed out of subject clitic doubling? In this case, too, the same assumptions will lead to the correct conclusion, with no theoretical elaboration. An instance of the sentence type that has typically been put forth to show the impossibility of extraction out of subject doubling in this dialect is given in (21b). For comparison, its declarative counterpart is provided in (21a).

- (21) a. Pierre il boit.
 Pierre he drinks.
 'Pierre drinks.' [Literally, 'Pierre he drinks.']
- b. *Qui il boit?
 who he drinks [expected reading: 'Who drinks?']

It has been assumed in the past that (21b) would be the result of extraction out of subject doubling in Pied Noir French. There is no reason, however, why the result of extraction should not lead to one of the structures given in (22), all of which are logically possible.

- (22) a. *Qui je bois? [expected reading: ‘Who drinks?']
 who I drink
- b. *Qui elle boit? [expected reading: ‘Who drinks?']
 who she drinks
- c. *Qui ils boivent? [expected reading: ‘Who drinks?']
 who they (masc) drink
- g. *Qui elles boivent? [expected reading: ‘Who drinks?']
 who they (fem) drink

The fact that the illicit sentence type given in (21) is the one that has typically been given in the literature on subject clitic doubling – not only in Pied Noir French, but also in other languages – is not fortuitous. It is, I believe, based on the tacit assumption that the third person masculine singular clitic is somehow unmarked. Based on a comparison of Berber and Pied Noir French, it is obvious that the clitic *il* does not have the same status as the neutral Berber clitic *y-n*. For if it did, sentences (19b) and (21b) should either both be grammatical or both be ungrammatical. This is not the case, however, as illustrated. Now if French *il* is treated as a third person masculine singular clitic, the overall paradigm is to be expected. The ill-formed question in (21b) is ruled out for exactly the same reason that the following structures are out: the starred French declaratives in (16), the Berber question in (19a), the Berber questions in (20), and the French questions in (22). In each case, there is a mismatch in the value for person, number, and gender features between a clitic and the DP it is coindexed with. As a consequence, extraction out of doubling constructions is impossible in Pied Noir French simply because there is no default clitic available.

5. The Range of the Extraction Phenomenon

We are now in a position to generalize the approach advocated to other instances of clitic doubling constructions, more particularly the clitic doubling paradigm of direct and indirect objects of Berber. The following examples illustrate extraction of direct objects in Taqbaylit, a dialect that allows optional clitic doubling of objects.

- (23) a. *Wala-gh amci.*
 saw-1s cat
 'I saw the cat.'
- b. *Acu_i ay wala-gh e_i?*
 what that saw-1s
 'What did I see?'
- c. *Wala-gh t_i wemci.*
 saw-1s him cat
 'I saw [him] the cat.'
- d. **Acu_i ay t_i wala-gh e_i?*
 what that him saw-1 [Literally, 'What did I see him?']

In (23a) and (23b) there is no clitic doubling. The extraction structure in (23b) poses no problem. The strings in (23c) and (23d) involve clitic doubling. The latter construction is presumably ruled out because of a prohibition against extraction out of clitic doubling structures. There is clearly a parallelism between these patterns and the subject doubling examples of Pied Noir French. In both cases doubling is optional. And in both cases there is no neutral clitic available. The same explanation that was provided for the impossibility of extraction out of subject doubling in Pied Noir French is applicable to the above data. The string in (23d) is ruled out because the object clitic *t* is the third masculine singular clitic, not a clitic with a neutral value. Since there is no default object clitic in Taqbaylit, the only source of extracted objects are representations that contain no clitic, ones that ultimately lead to patterns such as (23b). The same conclusion was arrived at with respect to extracted subjects in Pied Noir French.

The proposed analysis may be extended to the extraction of indirect objects. Relevant examples are provided in (24).

- (24) a. *Wci-x aysum i wqqzin.* 'I gave meat to the dog.'
 gave-1s meat to dog
- b. *Wci-x as_i aysum i wqqzin_i.* 'I gave [him] meat to the dog.'
 gave-1s him:dat meat to dog
- c. *Mumi_i wci-x aysum e_i?* 'Who did I give meat to?'
 whom gave-1s meat
- d. **Mumi_i as_i wci-x aysum e_i?*
 whom him:dat gave-1s meat
 [Literally, 'Whom did I give him the meat?']

Along the lines of the approach advocated, the unacceptability of (24d) is due to the presence of a clitic (namely *as*, the third person singular dative clitic) in the output of extraction of the indirect object DP. Since the trace in the clitic chain must have a default value for ϕ -features, while the clitic is marked for third person singular, the clitic chain is ill-formed. Given that no default dative clitic is available, the potential pattern in (24d) will never be generated. The only available *wh*-construction is (24c), again exactly as we saw with the extractability of Pied Noir French subjects and Taqbaylit objects. In all three cases, grammaticality goes hand in hand with the well-formedness of a clitic chain.

Before closing this section, a word should be said about LF-extraction. Clitic doubling constructions seem to resist quantifier raising as well. The question that naturally arises is whether the proposed analysis can be extended to cover those instances where quantifier raising is illicit out of doubling constructions. Such a proposal has in fact been made in the past. In Roberge (1986), for instance, a well-formedness condition together with a definition of a clitic chain are proposed to cover both syntactic extraction and LF extraction out of doubling constructions. In view of the facts examined in this paper, there is no reason to expect both syntactic extraction and LF-extraction out of doubling constructions to be handled by the same mechanism. And the reason is that syntactic extraction out of clitic doubling is rooted in the notion of agreement, while the scope of quantification, until proven otherwise, is not. This is a crucial difference which receives empirical support from the fact that the same cross-linguistic uniformity argued for above for syntactic extraction does not seem to hold at LF. The disparity between languages is particularly striking when very closely related dialects are examined. Compare in this respect Colloquial French and Pied Noir French, as discussed in Roberge (1986). Both dialects allow subject clitic doubling. The sentence in (25a) is grammatical in both dialects, but (25b) is acceptable in Pied Noir but not in Colloquial French.

- (25) a. Jean il aime manger.
 J. clitic like:3s eat:inf
 Jean likes to eat.
- b. Personne il aime manger.
 nobody clitic like:3s eat:inf
 Nobody likes to eat.

Even within the same language a lack of uniformity regarding LF-extraction manifests itself. In Berber for instance, where subject clitic doubling is

obligatory, a quantified DP subject may be headed by the quantifier *ijj* 'one', but not by the quantifier *kull* 'every'.

These discrepancies are a sufficient indication that the analysis proposed for syntactic extraction cannot be extended *mutatis mutandis* to cover Quantifier Raising out of clitic doubling structures. Syntactic extraction is rooted in a proper characterization of agreement, but LF-extraction is not. The LF clitic doubling phenomenon is perhaps best handled within the theory of binding in the manner suggested by Roberge (1986), where the options available are parametrized. In contrast, syntactic extraction calls for an agreement based analysis such as the one proposed in this paper.

6. Berber Dialects with no Subject Clitic Doubling

In a study devoted to the phenomenon of anti-agreement, Ouhalla (1993) provides an analysis of the anti-agreement effect observed in a number of languages, including Berber. Ouhalla does not treat the overt realization of Berber subjects as an instance of clitic doubling. His account of the anti-agreement effect is based on an adaptation of Aoun and Li's (1989) principle of A'-disjointness Requirement. According to his analysis anti-agreement amounts to a strategy utilized by some Null Subject languages to prevent a resumptive *pro* from appearing in certain subject positions, namely those that are accessible to binding by a moved *wh*-phrase. The reason for the prohibition is that agreement would lead to a violation of the A'-disjointness Requirement, a principle designed to regulate the distribution of pronominals. We will not go into the details of Ouhalla's proposal. The point that needs to be stressed, however, is that according to him the anti-agreement effect observed in Berber is not to be attributed to the interaction of clitic doubling and agreement. Such a view contrasts significantly with the one expressed in this paper. In what follows, I will present new evidence from dialects of Berber where anti-agreement is prohibited in constructions involving syntactic extraction of subjects. These dialects thus contrast with the ones examined so far. What is remarkable is that the proposed analysis predicts that the two dialect types will also contrast in other aspects of the grammar in a specific manner. As will be shown this prediction will be borne out by the facts to be examined.

In a discussion of the Siwa dialect of Berber, Laoust (1932) states that nominals in eastern Berber dialects, of which Siwa is an example, exhibit a unique morphological shape. His statement is as follows.

“Le nom berbère peut se présenter sous deux formes: a) une forme dite *absolue* sans modification de la voyelle préradicale; b) une forme dite d’*annexion* avec modification de cette même voyelle. Celle-ci se substitue à la première chaque fois que le nom est sujet du verbe et placé après lui; ou dépend d’un nom de nombre cardinal; ou est régi par certaines particules, prépositions pour le plus grand nombre. Mais ces règles, qui trouvent leur emploi constant dans les parlers algériens et marocains, sont inconnus des parlers de Tripolitaine et de Libye. Dans ceux du Djebel Nefousa, de Ghat, de Sokna, de Ghadamès, de Siwa, le nom s’offre, en effet, sous une forme *invariable* quelle que soit la particule qui le régit ou la place qu’il occupe dans la phrase.” (p. 96-97)

According to Laoust’s statement, Siwa nominals exhibit a unique morphological shape. They thus contrast with the dialects examined so far in this paper. More precisely, they do not exhibit the State alternation described in sections 1 and 2, a phenomenon Laoust refers to as *forme absolue* versus *forme d’annexion*. The Siwa nominals which appear in (26) exhibit the given shapes regardless of the syntactic context in which they appear.

(26) <i>masculine</i>		<i>feminine</i>	
ayddid	‘goatskin’	tasert	‘hand mill’
aseggas	‘year’	tadbirt	‘pigeon’
azemmur	‘olive tree’	tahmast	‘tooth’

Siwa Berber contrasts with western varieties of Berber such as the two varieties examined in the body this paper. In Ait Seghrouchen Berber, for instance, the equivalent forms of the above strings exhibit two different shapes depending on whether they are in the Construct State or in the Free state, as traditional terminology would have it.

(27)	<i>Free State</i>	<i>Construct State</i>
<i>masculine:</i>	ayddid	wyddid
	asggwas	wsggwas
	azemmur	wzemmur
<i>feminine:</i>	tasirt	tsirt
	tadbirt	ttbirt
	tighemst	tghemst

Taqbaylit forms are also similar to the examples in (27). Recall that the Construct State is a nominal involved in a syntactic construction where case marking appears not on the DP but rather on a clitic coindexed with the DP, or on the

case marker in a KP structure. The Free State, on the other hand, signals that case is marked directly on an DP, and nowhere else. Siwa nominals, examples of which appear in (26), exhibit a unique shape, a shape that corresponds uniquely to the Free State. All things being equal, we would expect case to be represented nowhere else except on DPs. What this means is that Siwa would exhibit neither clitic doubling nor overt case markers. What are taken to be case markers in Ait Seghrouchen and taqbaylit must be analyzed as prepositions in Siwa. One way of testing the conclusion that we have just drawn should come from the properties of syntactic extraction. More precisely, the analysis proposed in this paper makes the following prediction. There will be no anti-agreement effect in Siwa in subject extraction. This indeed is the case, as illustrated in (28) with two cleft-sentences from Laoust (1932, p. 119).

- (28) a. nic awggwid wen usi-gh itellin ghur-ek.
 me man that went-1s yesterday to-you
 'I am the man who went to your house yesterday.'
- b. nettatet talti ten t-us d.
 her woman that 3fs-went here
 'She is the woman that came.'

In (28a) agreement on the verb is the first person singular and corresponds to the first person clefted DP. In (28b) the agreement marker shown on the verb is third person feminine singular, again corresponding to the extracted DP *nettatet* 'her'. Siwa thus contrasts as expected with Ait Seghrouchen Berber, where the neutral clitic complex *y-n* is utilized regardless of the value for person, number, and gender of the extracted subject.

- (29) a. ntc d aryaz din gher-c y-rah-n idennat.
 me is man that to-you went yesterday
 'I am the man that went to your house yesterday.'
- b. nttatt d tamttutt din dd y-rah-n.
 her is woman that here went
 'She is the woman that came.'

The difference between Siwa and the other dialects is that in one case subject markers are agreement markers (Siwa), and in the other they are clitics. Agreement or lack of it in extraction contexts is dictated by the principles of the system proposed in this paper.

If Siwa subject markers are agreement markers, we would expect the verb in structures involving questioned subjects to exhibit the default third person masculine singular, not the neutral marker utilized in Ait Seghrouchen. This indeed is the case.

(30) <i>Siwa</i>	<i>Ait Seghrouchen</i>
bettin y-us d?	wi dd y-us-n?
who 3ms-came here	who here came
'Who came?'	'Who came?'

In this respect, Siwa behaves like Pied Noir French, Colloquial French, or Spanish. It utilizes default agreement in questions. We see then that those dialects that allow extraction with agreement are precisely the ones where there is no Construct contrast. And those that have a Construct contrast are the ones that do not allow agreement, exactly as expected.

7. Summary

The subject markers of the western Berber dialects have been shown to be clitics and not merely an expression of agreement. In addition, arguments in support of the view that the presence of overt lexical subjects in Berber are instances of clitic doubling have been provided. The evidence adduced in favor of this position comes primarily from those grammatical phenomena that have long been recognized to be associated with the clitic doubling paradigm. A uniform analysis of syntactic extraction out of doubling constructions is then proposed and generalized to all instances of clitic doubling. The approach advocated is based on the idea that clitics are to be base-generated under INFL. In some cases the presence of a clitic is obligatory (e.g. the western Berber subject clitic), in others it is optional (e.g. Pied Noir French clitics). The notion of clitic chain is introduced and defined as a pair consisting of a clitic and the DP it is coindexed with. The well-formedness of a syntactic string is concomitant with the well-formedness of a clitic chain. A clitic chain is well formed just in case there is a match between the features of its two members. A trace, which can be a member of a clitic chain, is assumed to have default value for agreement features. Consequently, when a trace is coindexed with a clitic, that clitic must also exhibit a default value for agreement features. In some languages, there is a phonetically realized clitic with default marking for such features (e.g. the Berber neutral subject clitic). In this instance, grammaticality goes hand in hand with the cooccurrence of the default clitic and the trace, in accordance with the proposed system. If any other clitic is utilized ungrammaticality results. In those languages where a default clitic for a particular grammatical function is not available, extraction is possible only when a clitic is not present at D-structure.

References

- AOUN, J. (1981) «The formal nature of anaphoric relations», Doctoral dissertation, M.I.T.
- AOUN, J. & A. LI (1989) «Two cases of logical relations: Bound pronouns and anaphoric relations», University of Southern California manuscript.
- BORER, H. (1983) *Parametric Syntax: Case Studies in Semitic and Romance Languages*, Foris, Dordrecht.
- CHOMSKY, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, Linguistic Inquiry Monograph 6, M.I.T. Press, Cambridge, MA, USA.
- GALAND, L. (1979) «Relations du verbe et du nom dans l'énoncé berbère», In *Relations Prédicat-Actants dans les langues de types divers*, C. Paris ed. LACITO documents, Eurasie 3, SELAF, Paris.
- GUERSSEL, M. (1983) «A phonological analysis of the construct state in Berber», *Linguistic Analysis*, vol. 11, p. 309-330.
- GUERSSEL, M. (1992) «On the case system of Berber», *Canadian Journal of Linguistics*, vol. 37, p. 175-195.
- JAEGGLI, O. (1981) *Topics in Romance Syntax*, Foris, Dordrecht.
- LAOUST, E. (1932) *Siwa: Son parler*, Ernest Leroux, Paris.
- MAY, R. (1985) *Logical Form: Its Structure and Derivation*, M.I.T. Press, Cambridge, MA, USA.
- OUHALLA, J. (1993) «Subject-extraction, negation and the anti-agreement effect», *Natural Language and Linguistic Theory*, vol. 11, p. 477-518.
- RIZZI, L. (1986a) «Null objects in Italian and the theory of *pro*» *Linguistic Inquiry*, vol. 17, p. 501-557.
- RIZZI, L. (1986b) «On the status of subject clitics in Romance», In *Proceedings of the XIV Linguistic Symposium on Romance Languages*, O. Jaeggli & C. Silva-Corvalan eds, Foris, Dordrecht.
- ROBERGE, Y. (1986) «On the syntactic recoverability of null arguments», Doctoral dissertation, University of British Columbia.