

The development of negation in Arabic and
Afro-Asiatic

Christopher Lucas

Emmanuel College

This thesis is submitted for the degree of

Doctor of Philosophy

at the University of Cambridge

November 2009

Abstract

This thesis discusses diachronic developments in the expression of negation in Arabic and other Afro-Asiatic languages, focussing in particular on the set of changes known as ‘Jespersen’s Cycle’ – prototypically the progression from preverbal to bipartite to postverbal negation – as well as the development of indefinites in the scope of negation. Drawing together data on negation from a number of neighbouring varieties of Arabic and Berber, as well as from Coptic and Modern South Arabian, this thesis defends from a linguistic and historical point of view the claim that bipartite negation in Arabic was triggered by contact with Coptic in Egypt, and separately with Modern South Arabian in Yemen and Oman, and that the same construction in Berber was in turn triggered by contact with Maghrebi Arabic. In light of the lack of an existing model of the psychological mechanisms which enable contact-induced grammatical change, as opposed to the sociolinguistic factors which constrain it, an account of these mechanisms is developed, integrating Van Coetsem’s (1988, 2000) work on this topic with research on second language acquisition and first language attrition, as well as with acquisitionist approaches to (internal) change in general. This then enables an explicit account of the spread of bipartite negation in the languages under study. This account sees the bipartite construction in Arabic as the product of imposition (source-language agentivity) by native speakers of Coptic and Modern South Arabian, and its counterpart in Berber as the result of borrowing (recipient-language agentivity) by native Berber speakers from their second-language Arabic. The partial and complex progression from a bipartite to a postverbal negative construction in Palestinian Arabic is then examined in detail on the basis of original field data, in a case study of phonological input to syntactic change. Finally, the scope is widened to investigate a number of Jespersen-type developments in the Semitic and Cushitic languages of Ethiopia, as well as the development of n-words and negative indefinites in Palestinian and Moroccan Arabic, Maltese and Hebrew, where it is argued that, contrary to initial impressions, only the latter two have developed into *bona fide* negative concord languages.

Acknowledgments

There are many people who deserve thanks for their direct or indirect involvement in making this thesis a reality. I will just mention the most important here.

First of all, I gratefully acknowledge the financial support of the Arts and Humanities Research Council.

Next, I want to thank my supervisor, David Willis. There is no doubt that I have been extremely fortunate in having David as my supervisor. It goes without saying that without his help, support and criticism this thesis would not have been possible: essentially everything I know about historical linguistics I have learnt from him. But more than for his excellent work as supervisor, I want to thank David for all his advice, encouragement and support, way beyond the call of duty, in the matter of making a career in linguistics. It remains to be seen whether this prospect is to become a reality as far as I'm concerned, but it is clear that the chances would be far smaller were it not for David's help.

I have also been very fortunate in being part of a larger project on the historical development of negation together with David and Anne Breitbarth. Being part of a very talented team working towards the same goal, and having another person who has taken a serious and constructive interest in my work, especially one as perceptive as Anne, has been a huge benefit.

Next I want to thank all the many people who have provided help and input for specific parts of this thesis. The most important of these is Elliott Lash. Elliott and I worked together on putting together and analysing the corpus of Coptic texts, and co-authored an article for the *Journal of Linguistics* partly on the basis of this corpus. This would never have been possible without Elliott's impressive Coptic skills, so thanks a lot for your help Elliott! All parts of the thesis which are the direct outcome of work done in collaboration with Elliott are clearly indicated in the text.

Other people who have helped to improve this thesis greatly with their comments and input (though of course any faults remain my sole responsibility) are the anonymous reviewers of various sections that I have submitted as journal articles, including one in particular who was later onymized as Jonathan Owens, and also Lameen Souag, Michael Spagnol, Hamid Ouali, Yaron Matras, Adam Ledgeway and all the Palestinians who contributed to, and helped me with, my fieldwork, in particular the Owaineh family.

Finally, I want to dedicate this thesis to my wife Antonia, who makes everything else worthwhile.

Declarations

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. The total word count, including footnotes, references and appendix, is 75,446 words.

List of abbreviations

ABS	absolutive	PTCP	participle
ACC	accusative	Q	question marker
AUX	auxiliary	REL	relative
CLEC	Common Lowland East Cushitic	SBJV	subjunctive
CMPR	comparative	TEMP	past temporal
COMP	complementizer		
CONSTR	construct state		
COP	copula		
DAT	dative		
DEM	demonstrative		
EMPH	emphatic particle		
EXPL	expletive		
F	feminine		
FUT	future		
IMP	imperative		
IMPF	imperfect		
INF	infinitive		
IRR	irrealis		
JUSS	jussive		
KSKO	Karnataka Saraswat Brahmin Konkani		
M	masculine		
MSA	Modern Standard Arabic		
NEG	negative		
NOM	nominative		
NONPST	non-past		
NPA	negative polarity adverb		
NPI	negative polarity item		
PL	plural		
PLD	primary linguistic data		
PRES	present		
PRF	perfect		
PROG	progressive		
PRT	particle		
PST	past		

List of tables

1.1	Jespersen's Cycle in the history of European languages	47
2.1	Percentages of stage II and stage III negative constructions in the Coptic corpus	71
2.2	Environments for <i>ən</i> -retention and dropping	75
5.1	Indefinites in Classical Arabic / MSA	203
5.2	Indefinites in Egyptian and Palestinian Arabic	214
5.3	Indefinites in Moroccan Arabic	222
5.4	Haspelmath and Caruana's (1996) Maltese <i>xi</i> - and <i>ebda</i> -series	223
5.5	Indefinites in Maltese	232
5.6	The evolution of Maltese indefinites	233
5.7	Distribution of <i>-š</i> without <i>ma</i> in Cairene	241
5.8	Grammaticality of negative constructions by context in Cairene	241
5.9	Acceptability judgments for stage III construction with imperfect	244
5.10	Acceptability judgments for stage II/III constructions with the perfect ...	245
5.11	Acceptability judgments for stage II/III constructions with labial-initial pseudo-verbs	246
5.12	Acceptability judgments for stage II/III constructions with pseudo-verb <i>šind</i> -	247
5.13	Acceptability judgments for stage II/III constructions with a labial-initial perfect verb.....	249
5.14	Acceptability judgments for question-marking <i>-š</i> with a perfect verb.....	251
5.15	Grammaticality of negative constructions by context in Palestinian.....	252
5.16	Tokens of negative constructions in the speech of young acquirers of Jordanian	271
5.17	Tokens of negative constructions in the speech of Jordanian mothers	272
6.1	Inflection of negative past tense and negative copula in Afar	291
6.2	Indefinites in Modern Hebrew	300

List of figures

1.1	Jespersen's Cycle in Europe	18
1.2	Jespersen's Cycle in Arabic	25
2.1	Jespersen's Cycle in Arabic and Berber	51
2.2	Three scenarios for the transfer of the stage II construction from Coptic to Arabic	80
4.1	How imposition leads to change	155
4.2	How borrowing leads to change	167
6.1	Ethiopian Semitic family tree	275
6.2	Cushitic family tree	276

Contents

1	Introduction	10
1.1	Afro-Asiatic	10
1.2	Jespersen's Cycle	13
1.2.1	Areal Jespersen's Cycle	17
1.3	Jespersen's Cycle in Afro-Asiatic	19
1.3.1	Arabic sentential negation	20
1.3.2	Coptic sentential negation	27
1.3.3	Berber sentential negation	31
1.4	Aims and outline of the thesis	34
2	Jespersen's Cycle in north Africa: the case for contact	38
2.1	Bringing contact in from the cold	38
2.2	How common is Jespersen's Cycle anyway?	43
2.3	The case for Arabic-Berber contact	50
2.3.1	Did stage II negation first arise in Arabic or Berber?	54
2.4	The case for Coptic-Arabic contact	65
2.4.1	Corpus study of negation in late Coptic	70
2.4.2	Further obstacles for the case for Coptic-Arabic contact	76
2.4.3	Jespersen's Cycle in the southern Arabian Peninsula	90
2.5	Postscript: Jerusalem Domari	99
3	Modelling contact-induced change.....	102
3.1	Outline	102
3.2	Problems with contact linguistics	103
3.3	Van Coetsem's model	106
3.4	Expanding and revising Van Coetsem's model	111
3.4.1	Dominance	111
3.4.2	The 'stability gradient'	123
3.4.2.1	Grammatical borrowing	126
3.4.2.2	Grammatical imposition	135
3.4.3	Restructuring	145
4	Mechanisms of contact-induced change	150
4.1	Coptic to Arabic imposition of stage II negation	152

4.2	Arabic to Berber borrowing of stage II negation	157
4.2.1	L1 performance attrition	157
4.2.2	Performance attrition in Berber and KSKo	160
4.3	English restructuring	167
4.4	Summing up	169
5	Focus on Arabic	172
5.1	Residual stage I and the semantics of stage II negation in Arabic	172
5.2	Developments in Arabic indefinites	186
5.2.1	Issues in the terminology and analysis of indefinites	188
5.2.2	N-words in Classical/MSA	196
5.2.3	N-words in Palestinian/Egyptian	204
5.2.4	N-words in Moroccan	215
5.2.5	N-words in Maltese	222
5.3	Stage III negation in Palestinian (and Cairene) Arabic	233
5.3.1	Previous literature	235
5.3.2	Research questions	241
5.3.3	Fieldwork methodology	242
5.3.4	Results and discussion	244
5.3.5	Reconstruction	252
5.3.5.1	NPI -š in Cairene	255
5.3.5.2	Stage III negation in Palestinian	263
5.3.6	Future prospects	270
6	Developments in the negation systems of other Afro-Asiatic languages ...	274
6.1	Previous literature	276
6.2	Jespersen-type developments in the Horn of Africa	282
6.2.1	Bipartite negation in Ethiopian Semitic	282
6.2.2	Bipartite negation in Lowland East Cushitic	288
6.3	Developments in indefinites	292
7	Conclusion	307
	Works cited	311
	List of References	313

1 Introduction

This thesis discusses diachronic developments in the expression of negation in Afro-Asiatic languages, with a primary focus on Arabic. In so doing, it aims to further our understanding of the mechanisms underlying grammatical change in general by exploring their role in detail in this restricted domain. In particular we will see that the occurrence in a number of neighbouring languages of the change with which this thesis is centrally concerned – the development of a bipartite negative construction – requires an explicit framework for understanding how language contact can lead to grammatical change. The elaboration of such a framework is the primary theoretical contribution of the present work.

In this introductory chapter I first provide an overview of the Afro-Asiatic family and the reasons for singling out the history of negation in Arabic for particular scrutiny (section 1.1). Most important among these is that Arabic (specifically the transition from Classical Arabic to some of the present-day colloquial varieties) provides us with a particularly clear and well-attested example of the set of changes that have come to be known as Jespersen's Cycle. An overview of Jespersen's Cycle and previous work on the topic (almost entirely restricted to European languages) is given in section 1.2, and the basic data from Arabic, Berber and Coptic are presented in section 1.3. An outline of the thesis in section 1.4 concludes the chapter.

1.1 Afro-Asiatic

Afro-Asiatic is among the largest and most heterogeneous of the widely accepted language families. It is generally sub-divided into six genera: Semitic, Berber,

Egyptian-Coptic, Cushitic, Chadic and Omotic,¹ comprising in excess of 300 languages between them. As such, we cannot hope to investigate the diachrony of negation in a fully representative sample of these languages at an interesting level of detail within the confines of the present work. Moreover, many of the languages of this family are critically underdescribed and few have a significant textual attestation dating back to before the period of European colonization (this is particularly true of Chadic and Omotic, which consequently feature only incidentally in this thesis). Languages that constitute significant exceptions to these generalizations are Egyptian-Coptic, a number of Semitic languages, and, to a lesser extent, some Berber languages. Even if we were to restrict ourselves just to these languages, however, the history of negation in each of them could not be done justice here. Hence my primary focus in the present work is on Arabic, and languages whose developments in the expression of negation are arguably linked to those of Arabic through contact (including Coptic, a range of Berber and Modern South Arabian languages, and, more incidentally, several Indo-European languages). Other Semitic, and certain Cushitic languages will be addressed to the extent that written records or comparative reconstruction allow us to observe, and offer explanations for, significant changes in the syntax of negation.

The focus on Arabic in particular is motivated primarily by practical considerations. First, a comparison of the present-day colloquial varieties with

¹ There is no consensus as to the correct sub-grouping of these genera, and the status of Omotic within Afro-Asiatic is the subject of particular controversy. Once thought to be a sub-group within Cushitic (Greenberg 1963), there is now serious doubt as to whether it really belongs to Afro-Asiatic at all (Newman 1980, Diakonoff 1996, Theil 2006). A compromise view is that Omotic was the first genus to split from Afro-Asiatic, such that the remaining genera together form a sub-group that has been labelled 'Erythraean' (Ehret 1995).

Classical Arabic shows that many of the former have undergone significant developments in the expression of sentential negation (Jespersen's Cycle; see section 1.2),² and developments in the polarity of vernacular Arabic indefinites, while less clear cut, raise some interesting theoretical questions. Second, the large number of well described modern varieties with minor but significant variations in the expression of negation allows for detailed and relatively secure comparative reconstruction of various areas of morphology and syntax (cf. Owens 2006) compared with many other Afro-Asiatic languages. This is necessary because of the near total lack of earlier written records of vernacular Arabic: (an approximation to) the language of the Qur'an became the written standard of the Arabic-speaking world from the earliest days of Islam and has remained so ever since, such that textual evidence of the development of the colloquial varieties is restricted to occasional minor deviations (chiefly phonological and lexical) from this standard.³ Finally, although there has been very little expressly historical work on negation in Arabic, it has attracted the interest of a number of synchronic linguists, both descriptive and theoretical, especially as regards the intersection with indefinites and negative polarity (e.g., Woidich 1968, Haspelmath and Caruana 1996, Benmamoun 1997,

² The same is not true of some of the other Semitic languages with long and extensive written histories, such as Hebrew, Aramaic or Akkadian. Discussion of these is mostly limited to section 6.3 on the development of indefinites and negative concord in Afro-Asiatic languages other than Arabic.

³ The language of literature from the seventh to ninth centuries (including the Qur'an) is generally referred to as Classical Arabic. The form of this language was rigidly maintained in almost all writing up to the late nineteenth or early twentieth centuries, from which time written Arabic tends to be referred to as Modern Standard Arabic. The differences between Classical and Modern Standard Arabic are chiefly lexical and stylistic, rather than morphological or syntactic: it is probably fair to say that no structure that is grammatical in Modern Standard Arabic is ungrammatical in Classical Arabic.

2000, Ouhalla 2002, Hoyt 2005, 2006). Thus we have a more solid basis for a detailed understanding of the mechanisms of language change that brought about the development of negation in Arabic than we do for any other Afro-Asiatic language.

1.2 Jespersen's Cycle

The historical development of negation is a topic which has become the focus of considerable interest in recent years, particularly in studies of European languages. For a sample see Mazzon (2004), Tieken-Boon van Ostade, Tottie and van der Wurff (1998), Van Kemenade (2000) and Wallage (2005) on English; Abraham (2003), Donhauser (1996, 1998) and Jäger (2008) on German; Burridge (1993), Breitbarth (2009) and van der Auwera and Neuckermans (2004) on Dutch and Low German; Catalani (2001), Detges and Waltereit (2002), Martineau and Mougeon (2003), Schwegler (1988) and Schwenter (2006) on French; Willis (to appear) on Welsh; Devos and van der Auwera (2009) on Bantu; and Beyer (2009) on West African languages. Bernini and Ramat (1992, 1996) offer a typological overview of negation across European languages, while Van Gelderen (2008) gives an overview of cyclical developments in negation from a number of different language families, and van der Auwera (2009, in press) provides a synthesis of observed developments and proposed analyses across a range of European and non-European languages.

A recurring feature of all of the aforementioned European languages, which appears first to have been discussed in detail by Jespersen (1917), and which has

been named for him since Dahl (1979), is Jespersen's Cycle.⁴ Jespersen himself described the pattern thus:

“The original negative adverb [in a given language] is first weakened, then found insufficient and therefore strengthened, generally through some additional word, and this in turn may be felt as the negative proper and may then in the course of time be subject to the same development as the original word.” (Jespersen 1917: 4)

This definition-cum-explanation of Jespersen's is not without its problems. For one thing, it is not clear whether it refers to the process by which individual negative morphemes become augmented over time (as in pre-Latin **ne + oenum* 'not one' > Latin *non* 'not'), or to the syntactic process by which a language goes from marking negation with one item (prototypically preverbal) to marking it with a distinct item (prototypically postverbal), via an intermediate stage in which both items co-occur. In general, when authors since Dahl (1979) have referred to Jespersen's Cycle, they have tended to focus on the latter, syntactic process. Jespersen (1917: 7-14) himself, however, seems to have seen both types of development as manifestations of the same general tendency; and the use of the term Jespersen's Cycle to refer to both of these processes simultaneously has some currency (e.g. Kiparsky and Condoravdi 2006). This thesis, however, is primarily concerned with the development of bipartite negation, and it is to this that the term Jespersen's Cycle will always be used to refer

⁴ It is mentioned in passing by earlier authors, such as Gardiner (1904) and Meillet (1912: 393-394); cf. van der Auwera (2009).

here. In languages which undergo Jespersen's Cycle, then, sentential negation at an initial stage is marked by a single preverbal element, as in the Old French example in (1).⁵ This we will call stage I.

(1) Jeo **ne** dis.

I NEG say.PRS.1SG

'I do not say.' (Old French)

At stage II some postverbal element, often formerly a noun of minimal quantity or a word meaning '(no)thing', is grammaticalized to form a bipartite negative construction together with the original preverbal marker, as in the contemporary standard French example in (2).

(2) Je **ne** dis **pas**.

I NEG say.PRS.1SG NEG

'I do not say.' (contemporary standard French)

⁵ It is an open question as to whether, for a given language to be said to have undergone Jespersen's Cycle, the position of the various negators with respect to the verb must necessarily follow the model of French (preverbal > 'embracing' > postverbal), or whether cyclic renewal is all that is required. West Germanic languages, for example, are widely agreed to have undergone Jespersen's Cycle, but, given their predominantly OV syntax, in various clause types negation in these languages will have been preverbal at all stages of the cycle. That said, main clauses with verb-second order in West Germanic do show the typical migration of negation from a preverbal to a postverbal position. In any event, the Afro-Asiatic languages that are the primary focus of this thesis follow the French model exactly, with stage I always being preverbal, stage II embracing and stage III postverbal.

At stage III the original preverbal negative element becomes optional and is then lost altogether, so that the innovative postverbal item now suffices as the sole, unmarked negator, as is arguably the case in contemporary colloquial French:⁶

(3) Je dis pas.

I say.PRS.1SG NEG

'I do not say.'

(contemporary colloquial French)

As van der Auwera (2009) points out, many authors depict Jespersen's Cycle in terms of a more complex scenario involving four or five stages, but the three-stage schematization adopted here has the advantage of clarity and simplicity. It should not be taken to imply that the constructions of each stage cannot co-occur in a given language: as we will see, they often do. Similarly, while it will sometimes be convenient to refer to a language as a whole as being at stage I/II/III, often we will see that a stage I/II/III construction is preferred or required in different syntactic domains in a single period of a given language. Note also that apparently all languages, whether they are at stage I, II or III, have items which can optionally co-occur with the ordinary means of expressing sentential negation for the purpose of polar emphasis. As such, it is important to be clear from the outset that, in the schematization adopted here, what distinguishes a language with a true stage II construction from a stage I language with the expected optional emphatic elements is

⁶ This highly simplified schematization of the French Jespersen's Cycle is due to Jespersen (1917: 7). For more in-depth treatments see the above references.

that the new negator in a stage II construction has to be capable in at least some contexts of being non-emphatic relative to its stage I counterpart.

A second problem with the content of the above quotation from Jespersen concerns his claim that 'original negative adverb is first weakened'. First of all, it is unclear whether we are to understand 'weakened' as referring to a loss of expressive force (semantic weakening) or a loss of phonetic material (phonological weakening), or, again, perhaps both. Secondly, there is no compelling reason to suppose *a priori* that the semantic or phonological weakening of the original preverbal negator is a necessary precondition for the development of a stage II construction. Indeed, we will see that neither kind of weakening appears to have been necessary or responsible for the development of stage II negation in Arabic.

1.2.1 Areal Jespersen's Cycle

One interesting feature of Jespersen's Cycle that has become clear as the data from studies of individual languages have been collated is the extent to which its distribution in the languages of Europe is clustered geographically (see Figure 1.1; cf. also Bernini and Ramat 1992, 1996).



Figure 1.1

Jespersen's Cycle in Europe

Traditional varieties spoken in the shaded areas have all undergone Jespersen's Cycle.

Progression from stage I to stage II (and in many cases from there also to stage III) has occurred in all the traditional varieties spoken within a contiguous area roughly indicated by the shading in Figure 1.1 (the situation in northern Italy is represented only approximately here; see chapter 2 for more details), but apparently not in other Indo-European languages spoken in Europe (including Irish, Scots Gaelic and many

Romance varieties of Spain and Italy). A clear implication of this geographical clustering is that the progression of a given language from stage I to stage II is strongly correlated with contact with another language which is itself also stage II (or conceivably stage III, but see section 2.4). It should be clear, however, that contact with a stage II language is neither a necessary, nor a sufficient condition for a stage I language to develop a stage II construction. It is not necessary because at least one of the above-mentioned languages must have been the first in this area to develop (independently) a stage II construction; and it is not sufficient because neither Romance varieties spoken south of northern Italy and the Pyrenees, nor German-influenced Slavonic varieties such as Sorbian and Czech have developed a stage II construction, despite having been in extensive contact with languages which had. This point will be important in chapters 2 and 3, when we consider the situation in Egypt and north Africa.

1.3 Jespersen's Cycle in Afro-Asiatic

It has not gone unnoticed that similar changes can be observed (or reconstructed) in the histories of various Afro-Asiatic languages. Gardiner (1904) discusses the rise of the postverbal negator *an* (< *iwn3*) in Late Egyptian, and more recently there has been work on this topic by Kickasola (1975), Meltzer (1990) and Winand (1997). The development of a bipartite negative construction in Arabic and Berber has been mentioned intermittently by, for example, Basset (1952), Brugnatelli (1986) and Chaker (1996), but appears not to have been the subject of detailed study before now. However, the basic facts of these three languages at least are relatively well known. They are presented in the following three subsections.

1.3.1 Arabic sentential negation

Classical Arabic had a variety of preverbal sentential negators, the oldest of which would appear to be *lā* (compare Biblical Hebrew *lo*), this being largely restricted to the negation of the imperfect verb forms:⁷

(4)	ʔakθaru-hum	lā	yaʔlamūna
	most.NOM-them	NEG	know.IMPF.3MPL
	‘Most of them do not know.’		(Qur’an 31:25 and <i>passim</i>)

In addition to this we find a negator exclusively for the past (*lam*) and one exclusively for the future (*lan*), as well as a negative copula *laysa*, and a rarely used general-purpose negator *ʔin*. A fourth, *mā*, could be used with both perfect and imperfect verb forms:

⁷ In common with many languages, the Arabic verbal system is neither exclusively tense-marking nor exclusively aspect-marking. Traditionally one talks of tenses, and ‘perfect’ and ‘imperfect’ are the usual labels given to the two non-periphrastic tense forms found in all varieties of Arabic. Perfect verb forms typically express reference to past time with perfective aspect, and they have exclusively suffixing person-number-gender morphology. Imperfect verb forms are mainly prefixing in their inflection and they typically express nonpast time or habitual/progressive aspect. However, the Classical Arabic negator *lam* always selects an imperfect verb (in what is traditionally called the ‘jussive’ mood), despite reference then being to the past.

- (5) wa-*mā* ḡalamū-nā wa-lākin kānū anfusa-hum
 and-NEG wrong.PRF.3MPL-us but be.PRF.3MPL selves.ACC-their
 ḡaḡlimūna
 wrong.IMPF.3MPL
 ‘And they did not wrong us but they wronged themselves.’ (Qur’an 2:57)

Mā as negator appears to be an innovation within the prehistory of Arabic, having its origin in the reanalysis of a homophonous interrogative pronoun which is also found in Classical Arabic, but which has largely been lost in the modern dialects (see the discussion of Faber 1991 in section 6.1). Lipínski (1997: §47.15) and, following him, Rubin (2005: 50), plausibly suggest that the bridging context for this reanalysis would have been rhetorical questions such as ‘what do I know?’ > ‘I know nothing’. We can imagine this reanalysis working as follows. By asking what are the members of a set picked out by a given predicate, when it is mutually manifest to the speaker and hearer that the speaker believes that this set is empty, a speaker is able in cases such as these to communicate her belief as an implicature, rather than as part of the literal content of her utterance.⁸ It is not hard to imagine, however, that frequent use of such a communicative strategy could result in the semanticization of the implicature and hence the reanalysis of a former interrogative pronoun as a negator.⁹

⁸ Here and throughout I adopt the arbitrary convention of making speakers female and hearers male.

⁹ Given an example along the lines of ‘what do I know?’ > ‘I know nothing’, one would expect an intermediate negative quantifier stage: interrogative > negative quantifier > negator. Since we have no textual evidence of the prehistory of this change, we have no way of knowing whether this was the case or not. Note, however, that another possibility is that the reanalysis took place in contexts where *mā* was the pseudo-argument of a verb such as ‘to profit’, ‘to

Precisely what factors provoked this reanalysis in Arabic in particular, when presumably such implicatures are available in all languages, remains unclear however.

The negators other than *mā* are becoming obsolete in the modern Arabic dialects. Those dialects which have remained at stage I of Jespersen's Cycle retain the classical negative construction with preverbal *mā*.¹⁰

(6) ʔal-lon *mā-ḥabbēt-ha*
 say.PRF.3MSG-to.them NEG-love.PRF.1SG-her

'He told them, "I didn't fall in love with her."' (Syrian Arabic; Brustad 2000: 284)

A stage II construction, however, is (generally speaking) the unmarked structure for expressing sentential negation in non-Bedouin¹¹ Arabic varieties spoken across coastal north Africa (Chaker and Caubet 1996; Woidich 1968) and Upper Egypt (Khalafallah 1969), in Palestine, southern Lebanon and parts of Jordan (Shahin

succeed' etc., which quite often feature an expression of the extent of profit/success that is potentially ambiguous between an adjunct and an argument, e.g. 'What does it profit a man, if...?'. A former interrogative pronoun could, in such contexts, still (immediately) be analysed as a non-argument (i.e. a negator) rather than an argument (i.e. a negative quantifier) (cf. Breitbarth, Lucas and Willis forthcoming).

¹⁰ Some dialects have retained the long *ā* in this item, others have shortened it. Outside of examples from specific dialects I use the spelling *mā* to refer to the item crossdialectally.

¹¹ A reasonable generalization in Arabic dialectology is that a Bedouin dialect will tend to have more in common with another spoken in a different region than it does with the sedentary dialects of the same region. One typical feature of Bedouin dialects is conservatism in the expression of negation (as well as quite generally). There are of course individual exceptions to both of these generalizations, particularly in cases where Bedouin have become sedentarized and (partially) integrated into sedentary communities, when dialect hybridization tends to result.

2000; Palva 2004), in parts of Yemen (Watson 1993; Behnstedt 1985), in parts of Oman (Reinhardt 1894), and in Malta (Borg and Azzopardi-Alexander 1997). All other dialects appear to have remained at stage I (see Figure 1.2). The Arabic stage II construction is composed of the same preverbal *mā* (sometimes with a reduced vowel),¹² together with a postverbal enclitic which has been grammaticalized from the word for '(any)thing': *šayʔ* in Classical Arabic, usually *šī* in those modern dialects where this form retains its original meaning. As a negator, this form has generally become an enclitic *-š*, as in (7) from Cairene, (8) from Ṣanʿāni (Yemen) and (9) from Moroccan:¹³

- (7) **ma**-bəḥibb-**š** migiyy-u hina ktīr
 NEG-like.IMPF.1SG-NEG coming-his here much
 'I don't like his coming here a lot.' (Cairene; Woidich 1968: 33)
- (8) bih nās **mā** yiʿjib-hum-**š** aš-šāy
 there.is people NEG please.IMPF.3MSG-them-NEG the-tea
 'There are people who don't like tea.' (Ṣanʿāni; Watson 1993: 261)
- (9) **ma** nkdəb-**š** ʿli-k
 NEG lie.IMPF.1SG-NEG on-you
 'I'm not lying to you.' (Casablanca Moroccan; Adila 1996: 103)

¹² I will not take a position here on whether, and in which varieties, *mā* has become a clitic on the verb, except in the clearest cases. Where examples are taken from grammars I simply follow the conventions that each grammar adopts as regards writing *mā* with a hyphen attaching it to the verb or not. Elsewhere I write it as a separate word.

¹³ Heath (2002: 212) notes that an unreduced form *-šī* is common in northern Moroccan dialects, while Khalafallah (1969: 100-102) gives *-šey* as the ordinary form for the variety of Ṣaʿīdi (southern) Egyptian that he describes.

A related development is that of a constituent negator and negator of non-verbal (including participial) sentences *muš ~ miš* (10) from a reduced form of the third person masculine singular negative copula *mā-hū-š*.

(10) šuft ḥāga **miš** mafʔūla
 see.PRF.1SG thing NEG thinkable

‘I saw something unbelievable.’ (Cairene Arabic; Woidich 2006: 338)

The negative copula itself is composed of a pronoun and negation marking (11), and presumably arose through reanalysis of a (negated) resumptive pronoun following a left-dislocated subject (12) (a common source of copulas; cf. Pustet 2003: 55-56).

(11) il-ʕumra **ma-hiyyā-š** farḍ il-ʕumra sunna
 the-umra NEG-it.F-NEG obligation the-umra custom

‘The *umra* (minor pilgrimage) is not a religious duty, the *umra* is *sunna* (following the customs of the Prophet).’

(Egyptian; Brustad 2000: 298)

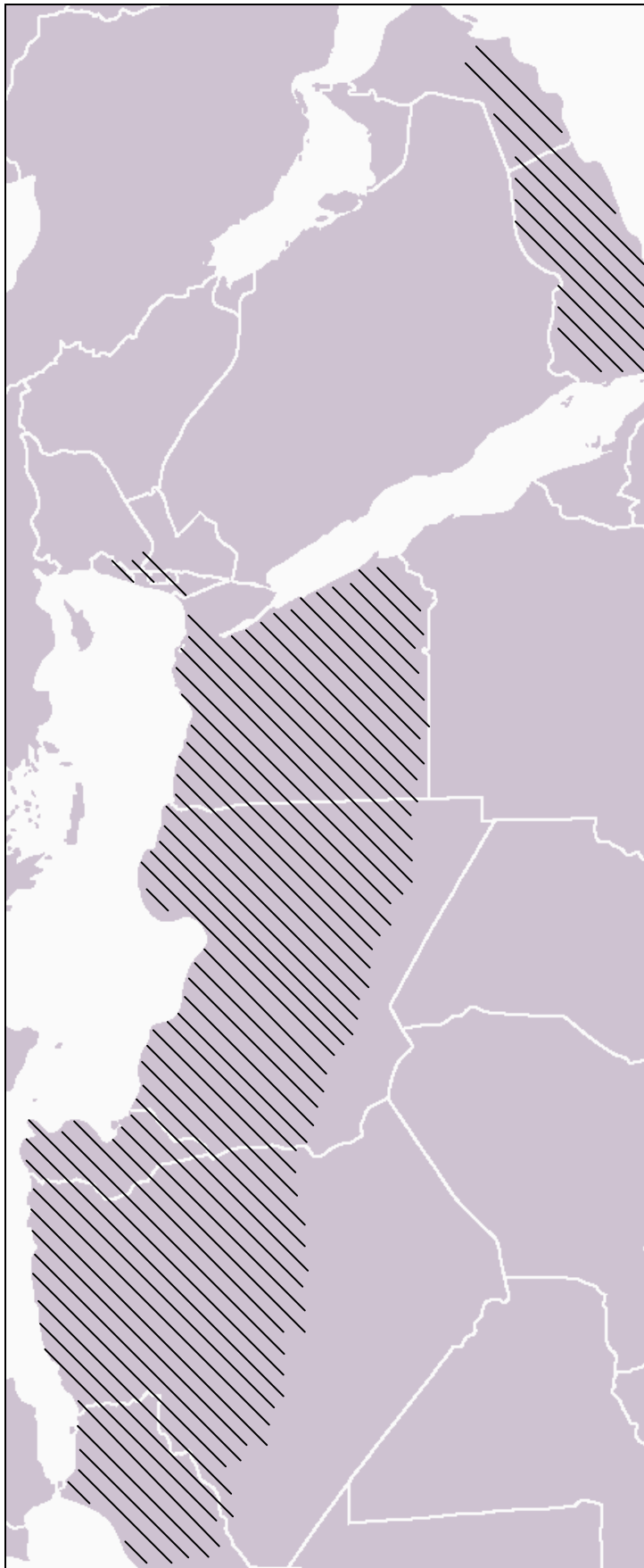


Figure 1.2

Jespersen's Cycle in Arabic (cf. Obler 1975: 33)

Shaded areas indicate regions in which a stage II/III Arabic variety is currently spoken. All other regions stage I or non-Arabic-speaking.

(12) aḥmad mā-hū-š ḡabi

Ahmad NEG-he-NEG stupid

Conservative interpretation: 'Ahmad, he's not stupid.'

Innovative interpretation: 'Ahmad isn't stupid.'

Dialects which have not undergone Jespersen's Cycle tend to have a negative copula composed of *mā* plus a pronoun without enclitic *-š*, and a corresponding constituent negator *mū* (not *muš*).

Among the dialects with a stage II construction there are several that also have an optional stage III construction. The best attested is Palestinian (Obler 1975, Palva 1984, 2004, Shahin 2000), where it seems that, in those circumstances where it is grammatical (to be discussed in detail in section 5.3), the stage III construction is in fact the unmarked structure, for example:

(13) ana (mā) bašrab-š il-ʔahwa

I (NEG) drink.IMPF.1SG-NEG the-coffee

'I don't drink coffee.'

(Palestinian Arabic)

A stage III construction is also reported for the southern Lebanese dialect of Baskinta (Abu Haidar 1979), the Jordanian dialect of es-Salt (Palva 2004), the Ṣaʿīdī dialect of Upper Egypt (Khalafallah: 1969), and the Omani dialect described by Reinhardt (1894). Further research on the latter dialect (if it still exists) would be particularly welcome as it appears to be unique in having a stage III negative copula (i.e. without preverbal *mā*):

(14) huwa-šī sekrān

he-NEG drunk

'He isn't drunk.'

(Omani; Reinhardt 1894: 282)

1.3.2 Coptic sentential negation¹⁴

Coptic is the latest stage of the Ancient Egyptian language, and shares many features with its Late Egyptian predecessor Demotic, being distinguished from the latter chiefly by its Greek-derived script. Coptic texts are found from the first century onwards, but Coptic is thought to have died out as a spoken language by the sixteenth century at the latest (Watterson 1988). However, it continues to be used in the liturgy of the Coptic Church until today.

Sentential negation in Coptic is accomplished through a complex set of constructions that are sensitive to various factors such as tense, modality, predicate type and predicate definiteness (in the case of nominal predicates). Only a brief overview of this intricate system will be given here.

Coptic has SVO basic word order and is predominantly prefixing (rather than suffixing). Verbal constructions usually have AuxSVO order, with the auxiliary expressed as a proclitic on the subject.¹⁵ The subject may be either a pronominal prefix or a full nominal expression, as shown in (15)a,b:

¹⁴ The data on negation in Coptic presented in this section (and in section 2.4.1) were collected and chosen in collaboration with Elliott Lash and are to be published in a forthcoming jointly authored paper (Lucas and Lash forthcoming).

¹⁵ One exception to this is the future tense, which is normally formed by prefixing *na-* directly to the verb.

(15) (a) a-i-cine əm pa-eiōt
 PRF-1SG-find ACC my-father
 ‘I found my father.’

(b) nere-**p-rōme** kōt ən ou-ēi
 IMPF-the-man build ACC a-house
 ‘The man was building a house.’

In negative contexts, two main strategies are used. Either a negative particle (simply encoding negation) is added to the auxiliary/verb, which otherwise remains as it would be in non-negative contexts, or a special negative auxiliary (encoding other features such as tense in addition to negation) replaces the affirmative auxiliary. The negative particle strategy is used with the present, imperfect and future tenses as well as certain constructions having causative or various temporal/conjunctive meanings (i.e. ‘until’, ‘then’, ‘when’, ‘and so’ + verb) and in several types of complement clause. The negative auxiliary strategy is found replacing the habitual, optative and perfect affirmative auxiliaries. The following examples illustrate the range of possibilities available for negation. The auxiliary strategy is illustrated in (16)a,b, which demonstrate the affirmative perfect auxiliary, *a*, and the negative perfect auxiliary, *əmpə*, respectively.

(16) (a) a-f-ouōšt na-i
 PRF-3MSG-worship DAT-me
 ‘He worshipped me.’

Although the *ən...an* form of the particle strategy shown in (18)a,b is only one of the many different negative constructions found in Coptic, there is a case for saying that it is the basic, or unmarked, construction, in that it alone is used in the negation of the present, imperfect and future tenses, as well as with clauses of attendant circumstance, and certain of the so-called ‘second tenses’, which place strong emphasis (or perhaps more accurately, focus) on some element in the sentence (other than the verb), frequently an argument or adverbial element (Lambdin 1982: 52, Reintges 2004). It is also the only option in the negation of verbless sentences (Loprieno 1995: 141):

- (19) **ən** anon ən-šēre ən-t-həmhal **an**
 NEG we the-children of-the-slave.woman NEG
 ‘We are not the children of the slave woman.’ (Gal 4:31)

Finally, Coptic also features a stage III construction alongside the bipartite *ən...an* construction as in (20), where the focus of negation – the clefted subject *anon* ‘we’ – is not preceded by *ən* (Loprieno 1995: 141):

- (20) anon **an** pe-nt-a-n-mere p-noute
 we NEG COP-REL-PRF-1PL-love the-God
 ‘It is not we who loved God...’ (I John 4:10)

1.3.3 Berber sentential negation

Berber languages are spoken principally in Morocco and Algeria, but also in Tunisia, Libya, Egypt, Mauritania, Mali, Niger and Burkina Faso (as well as in numerous expatriate communities around the world). Negation in these languages presents the following picture. Almost all varieties have in common that negated verbs undergo a form of ablaut which raises the stem vowel to /i/ (/e/ in Tuareg).¹⁶ Varieties which are conservative with respect to the expression of negation have a stage I construction with a single preverbal marker *ur* (or one of its cognate manifestations *wer*, *war*, *ul*, or *u*):

- (21) *ur* *igle*
 NEG leave.PRF.IRR.3MSG
 ‘He didn’t leave.’ (Tuareg; Chaker 1996: 16)

The most prominent representatives of this conservative negative construction are Tuareg (spoken in southern Algeria and southern Libya, Niger, Mali and Burkina Faso) and Tashelhiyt (spoken in southern Morocco). Note that these are regions in which either Arabic is not spoken, or the variety spoken is Ḥassaniyya, which has only the stage I construction exemplified (by Syrian Arabic) in (6).

¹⁶ We will not be concerned with this feature of Berber negation elsewhere in the thesis, but it is worth noting here, following Basset (1952: 15), Picard (1957) and Chaker (1996: 18), that this process of ablaut, which is usually referred to as the ‘negative theme’, was most likely not originally restricted to negative contexts. This remains the case in Kabyle, where verbs also undergo this process in other contexts such as conditionals (Chaker 1996: 18). By contrast a number of Tashelhiyt varieties lack this process altogether (Chaker 1996: 17).

In other varieties, most prominently represented by Central Atlas Tamazight (central Morocco) and Tarifit (northern Morocco), there is a stage II construction, of which the second element is *ša* in Central Atlas Tamazight and *ša* or *ši* in Tarifit. Both of these forms are cognate and can be traced back via a series of regular sound changes to an item *kra* ‘thing’ (Brugnatelli 1986, Chaker 1996: 16). The phonological resemblance to Arabic *-š* thus appears to be coincidental. Although *ša* in Central Atlas Tamazight can still function as an argument ‘thing’, at least in the scope of negation (Boumalk 1996: 41), it has undoubtedly developed a non-argumental function as a pure, non-emphatic negator. The same is true of Tarifit *ši*:

- (22) **ur** *iffiy* **ša**
 NEG exit.PRF.IRR.3MSG NEG
 ‘He didn’t go out.’ (Central Atlas Tamazight; Boumalk 1996: 36)
- (23) **ur** *izri* **ši** *imma-s*
 NEG see.PRF.IRR.3MSG **NEG** mother-his
 ‘He hasn’t seen his mother.’ (Tarifit; Boumalk 1996: 36)

Nevertheless, the second element appears always to be optional in these varieties. This contrasts with the situation in the two major varieties of northern Algeria, Kabyle and Shawia, where the second element (*-š(a)* in Shawia and a range of forms in

different dialects of Kabyle, including *kra*, *ara* and *ani*) is obligatory (24), except in certain well-defined contexts, to be discussed in sections 2.3 and 5.1.^{17, 18}

- (24) *ul ittaggad *(kra)*
 NEG fear.IMPF.3MSG NEG
 ‘He is not afraid.’ (Kabyle; Rabhi 1996: 25)

Finally, there are at least three Berber languages, Sened (Tunisia; Provotelle 1911), Aujila (Libya; Paradisi 1961) and Ghadames (the border of Libya, Tunisia and Algeria; Motylinski 1904, Mettouchi 1996), all of which are endangered or extinct, in which the postverbal element is an obligatory verbal enclitic *-ka/-č*, and the preverbal element is optional or entirely absent.

- (25) *akellim iššen-ka amakan w-iššen-ka tebārut*
 servant know.PRF.3MSG-NEG place know.PRF.3MSG-NEG road
 ‘The servant didn’t know either the area or the road.’
 (Aujila; Paradisi 1961: 82)

¹⁷ Shawia *š(a)* is straightforwardly derived from *kra* ‘thing’, as, trivially, is Kabyle *kra*. Chaker (1996: 16) derives Kabyle *ara* from a separate item *wara* ‘anything’, while Mettouchi (1996: 191) derives it from *haret* ‘thing’. It is unclear what the provenance of the form *ani* is.

¹⁸ An asterisk outside parentheses indicates that the item contained within them must be present in order for the sentence to be grammatical. An asterisk in front of an item within parentheses indicates that the sentence in question is only grammatical if that item is absent.

1.4 Aims and outline of the thesis

Taken together, the data on negation in Arabic, Coptic and Berber are suggestive of a spread of stage II negation between these languages as a result of contact. Similarly to what we observed in section 1.2 for European languages, there seems to be a correlation between progression from stage I to stage II in a given variety of one of these languages and contact with another that is itself also stage II. Of course, it is possible that this areal distribution is coincidental and the presence of stage II (or stage III) negative constructions in the relevant varieties of these three languages is the result of independent parallel developments. The purpose of chapter 2 is to assess the case for Jespersen's Cycle in Arabic, Coptic and Berber being related by language contact. The conclusion will be that, while we cannot definitively rule out the possibility that the occurrence of Jespersen's Cycle in Coptic and certain varieties of Arabic and Berber were entirely internal developments, nevertheless a strong case can be made for contact having played a crucial role.

This kind of caution in formulating the assessment of the situation is necessary, given the rather widespread preference for seeking internal explanations of linguistic changes over those based on contact. What the basis of this preference might be is explored in section 2.1, while the related notion that Jespersen's Cycle is too common a development to warrant positing a role for contact in its occurrence in a particular language is addressed in section 2.2. Sections 2.3 and 2.4 then make the case for Arabic-Berber and Coptic-Arabic contact, respectively, in the occurrence of Jespersen's Cycle in these languages.

It seems likely that at least part of what underlies the well-established skepticism concerning contact-based explanations of syntactic change is the fact that internal

mechanisms of change are simply much better worked out than their external counterparts. Moreover, all too often those who do not share this skepticism are liable to invoke contact as a brute, unexplained force, a kind of *deus ex machina* capable of almost anything. (For a randomly selected example, see the claim of contact with Semitic in the formation of Germanic verbal ablaut in Mailhammer 2007).

Chapter 3 seeks to address this imbalance by first reviewing some of the problems that characterize much existing work on language contact, and then outlining a more explicit model for how language contact can lead to syntactic change, taking as its starting point Van Coetsem's (1988, 2000) distinction between source-language and recipient-language agentivity, and adapting his insights in the light of research on second language acquisition and first language attrition.

One of the respects in which the model developed here diverges from Van Coetsem's approach is in subscribing to a key assumption of the approach to historical syntax adopted by generativists such as Lightfoot (1989, 1991, 1997, etc.) and Roberts (1994, 2001, etc.), but also non-generativists such as Harris and Campbell (1995). This is the innatist principle that the ultimate object of study in diachronic syntactic research is change in the mental grammars (I-language) acquired by individual speakers, rather than merely in their performance (E-language).¹⁹ That is, the focus is on speakers' knowledge of language and how this apparently stable object can come to change over time, rather than on their use of this knowledge, whose inherent instability is the product of irreducibly many social and psychological factors. However, like Harris and Campbell, but unlike Lightfoot

¹⁹ Though performance is, of course, the primary means we have for studying competence, particularly when the speakers in question are long dead.

and Roberts, I avoid hitching the model presented in chapter 3 to a particular syntactic formalism (such as mainstream generative grammar). This is partly because of the importance to the model of insights from outside formal syntax, but primarily because it aims to be sufficiently general as to be of use to linguists working within any, or no formalism, provided they can accept a distinction between knowledge of language and use of that knowledge.

With the model presented in outline in chapter 3, chapter 4 then makes detailed suggestions for some of the primary mechanisms by which language contact can lead to syntactic change, illustrating these mechanisms by means, among others, of the case studies of Coptic-Arabic and Arabic-Berber contact in the development of stage II negation.

Chapters 2 to 4, then, deal in relatively broad terms with how contact could have led to a major shift in the basic means of expressing negation in a number of north African languages. Chapter 5, by contrast, focusses more narrowly on the details of the evolution of negation in Arabic.

Section 5.1 offers a diachronic explanation for a number of syntactic restrictions on the stage II construction crossdialectally, and discusses the semantic status of the two elements of the bipartite construction, given that in some varieties either one or both of the two can be used interchangeably to express a single logical negation.

This sets up the discussion in section 5.2 of the question of which, if any, Arabic varieties exhibit 'negative concord', that is, the co-occurrence of sentential negation and one or more negative indefinite pronouns, such that the sentence is, again, interpreted as containing a single logical negation.

Finally, section 5.3 deals with the progression from stage II to stage III in Palestinian and, less clearly, Cairene Arabic, offering an internal reconstruction that explains both what precipitated the development of purely postverbal negation with *-š* in each variety, as well as the rather idiosyncratic restrictions on its distribution.

In chapter 6 the scope is then widened once more, as the methods and principles developed in previous chapters are applied to aspects of the development of negation in a number of other Afro-Asiatic languages. Several Jespersen-type developments in the Semitic and Cushitic languages of the Horn of Africa are analysed, and we investigate the rather limited development of indefinites in the scope of negation in Afro-Asiatic languages other than Arabic, most notably Hebrew.

Chapter 7 concludes the thesis.

2 Jespersen's Cycle in north Africa: the case for contact

This chapter begins by discussing in section 2.1 the traditional mistrust of contact-based explanations of linguistic change, and gives some reasons for thinking that an *a priori* preference for internal over contact-based accounts of change is unwarranted. In section 2.2 we then look at how these issues apply to the case of Jespersen's Cycle in general, and I assess the question of whether Jespersen's Cycle is too common a change to ever warrant an account involving contact in an individual instance. I argue against this view, and, on the basis of these arguments, section 2.3 makes the case for Jespersen's Cycle in Berber having been triggered by contact with Arabic. Section 2.4 then makes a similar case for contact with Coptic in triggering the Arabic Jespersen's Cycle, drawing data for Coptic from a corpus study of ninth and tenth century texts. Section 2.5 completes the picture of Jespersen's Cycle as a north African areal phenomenon with a brief discussion of bipartite negation in the Indo-Aryan language Jerusalem Domari.

2.1 Bringing contact in from the cold

The study of linguistic change has traditionally marginalized the role of language contact in this process, or sought to impose stringent theoretical constraints on its potential extent. What is more, this trend has continued almost unabated into recent times. Sapir (1921: 206), for example, held that all the evidence points to contact-induced morphosyntactic changes being “but superficial additions on the morphological kernel of the language,” and it probably remains uncontroversial in current historical linguistics to say that “an endogenous [i.e., purely internal – CL]

explanation of a phenomenon is more parsimonious [than one invoking contact – CL], because endogenous change *must* occur in any case, whereas borrowing is never necessary” (Lass 1997: 209). At the same time, mainstream generative and other approaches to the architecture of the human language faculty have tended to view this faculty as monolingual in essence, abstracting away from instances of bilingualism as a supposedly necessary idealization.²⁰ Hence Chomsky’s (1985) famous remark:

“Why do chemists study H₂O and not the stuff that you get out of the Charles River? ... You assume that anything as complicated as what is in the Charles River will only be understandable, if at all, on the basis of discovery of the fundamental principles that determine the nature of all matter, and those you have to learn about by studying pure cases.”

From this perspective monolingualism is the pure case while bilingualism, some degree of which is clearly a necessary condition for contact-induced change to occur, is merely a distracting pollutant.

While there was a time when these positions seemed unexceptionable, a gradually accumulating mass of evidence and argument to the contrary is starting to bring them into question. Some particularly striking recent evidence for the crucial role played by contact in determining the typological features of many of the world’s

²⁰ Following common practice, I use bilingualism here and throughout to refer to the state of having some knowledge of one or more languages in addition to one’s first, or native, language. This use of the term encompasses, but is not restricted to, cases of ‘balanced’ bilingualism, where individuals possess two or more first languages.

languages comes from the *World Atlas of Language Structures* (WALS; Haspelmath et al. 2008). Many of the maps in this important resource make clear the extent to which languages with the same features cluster together in the same geographical area, despite often belonging to quite separate genera and families.²¹ Of course, one should always consider the possibility that a typological trait shared by neighbouring languages arose in each of them as a result of independent internal developments. However, it is not clear that the widespread practice of ruling out on principle a role for contact in a given change, where a similar change can be shown to have occurred elsewhere purely as a result of internal development, is theoretically sound. It has often been assumed (though rarely explicitly stated) that factors involved in internal and external causation of diachronic change are necessarily mutually exclusive, but this is likely to be a false dichotomy. As Thomason and Kaufman (1988: 59) point out:

“it is no more reasonable to extrapolate a particular internal motivation from one case to another than it would be to extrapolate an external motivation from one case to another... If a reasonable external explanation for a change is available, it must not be rejected merely because similar changes have occurred under different antecedent conditions.”

For any given change that is accepted to have involved contact, there will be a range of internal factors that have combined with the external input to result in the change –

²¹ See especially, e.g. “Glottalized consonants” (Maddieson 2008), “Predicative possession” (Stassen 2008a), and “Comparative constructions” (Stassen 2008b).

it would make very little sense to say that internal characteristics of the language under study played no role in a change resulting from contact. But, by the same token, there is no justification for excluding contact on principle as a possible contributing factor in a change which one might also expect to arise purely internally. Whether one can actually *prove* that a certain development is due to contact seems doubtful, but, again, the same applies to claims of purely internal change. Loanwords are, of course, an exception to this uncertainty; but where syntactic change is concerned, the best we can usually hope for is sufficient evidence to say that a particular change is *likely* to have arisen through contact or through purely internal factors (cf. Heine and Kuteva 2007).

The extent to which theoretical opinion continues to diminish the importance of contact-induced syntactic change should not be overstated, however. Indeed, the whole notion of 'internally-caused' change has been called into question recently, most notably in the form of the 'Inertia Principle' (Keenan 2002, Longobardi 2001), which suggests that syntactic change should not occur in the absence of a syntax-external cause. This formulation does not exclude the possibility of syntactic change that is not the result of contact, but is triggered through the interface of syntax with phonology and the conceptual-intentional system. More extreme formulations of this idea have, however, been mooted (e.g. Meisel 2008), to the effect that there is no syntactic change without language contact. While this may be taking things too far (cf. Fuß, Roberts and Trips 2009), the rise of these kinds of views, particularly within generative circles, underlines the necessity of developing explicit models of how language contact can trigger syntactic change. It also shows the need for these models to be integrated with existing innatist approaches to change (such as that of

Lightfoot, Roberts and others), which have tended thus far to focus exclusively on language-internal triggers (but see Roberts 2007: ch. 5).

Regarding the centrality of bilingualism to the study of human language, it is probably still worth stressing first of all that it is by no means as uncommon worldwide as one might expect on the basis of the situation in Britain or the USA, for example. Cook (2002: 2) even goes as far as to claim that “[a]rguably the majority of people in the world are multi-competent users of two or more languages rather than mono-competent speakers of one language, and there are as many children brought up with two languages as with one.” Of course, it is entirely possible that the human language faculty is put to uses other than those for which it evolved (cf. Chomsky’s (2000) playing down of communication as the primary ‘purpose’ of this faculty). However, recent work on child bilingual acquisition has led researchers such as Meisel (2001: 40-41) to argue that

“the human language faculty predisposes the individual to become multilingual... [A]n adequate theory of grammar and of grammatical development must be capable of explaining multilingual development as the simultaneous acquisition of two or more first languages, i.e., an achievement of the human mind for which monolingual development is just a special case.”

Thus there is growing recognition of the centrality of bilingualism in the operation of the human language faculty and the need to integrate an account of the (bilingual) cognitive processes leading to contact-induced change into an adequate theory of

the nature of grammatical change in general. It is hoped that this thesis will go some way towards achieving these aims.

2.2 How common is Jespersen's Cycle anyway?

None of the above should be taken to imply, however, that the onus is no longer on the proponent of a contact-based account of a given change to make a strong case for that account. The intended conclusion is merely that an account involving contact should not be automatically rejected as soon as a reasonably satisfying purely internal account can be found. Each case should be addressed on its own merits and the evidence for and against contact assessed accordingly.

One of the most important stumbling blocks for any contact-based account of the development of bipartite negation in a given language is a widespread perception that Jespersen's Cycle is extremely common, which is then taken to automatically imply that any appeal to contact is spurious, in keeping with the sentiment of the Lass (1997: 209) quote in the previous section. Even if we dispute the validity of this logic, it is still worthwhile investigating the initial premise.

A recent explicit statement of this view comes from Van Gelderen (2008: 195), who claims that "cross-linguistically, the Negative Cycle may be one of the most pervasive of cyclical changes". In a footnote, however, she points out that, of the 1011 languages sampled in Dryer's (2008) *WALS* map on 'Negative morphemes', only 66 are listed as having bipartite negation, recognizing that these data appear to run counter to her claim. She goes on to acknowledge that "the use of multiple negative forms to express negation should be more common (e.g. standard French *ne pas*) if languages are continually undergoing negative cycles".

Her response at this point is to take issue with the data, rather than to reassess her generalization. She appears to accept the justification of the article accompanying Dryer's map for assigning to French a 'negative particle' rather than 'double negation' (= bipartite/stage II negation), since preverbal *ne* is optional in colloquial French. However, she complains that various languages that she investigates in her article are wrongly classified on the map as having single rather than bipartite negation. One set of languages that she mentions in this connection are the various Berber languages she gives data for. However, in her section on these languages she states (2008: 228), following Ouali (2003: 3), that "all (Northern) Berber dialects have a preverbal negative element *ur/wer* and most have an *optional* post-verbal element" (emphasis added). It is not my intention here to defend the accuracy of the data in *WALS*, nor to assess the usefulness of Dryer's map on 'Negative morphemes' in deciding to what extent Jespersen's Cycle or negative renewal is a 'universal'. However, it must be admitted that Dryer's not assigning any of the Berber languages in his sample to the category of bipartite negation is an accurate reflection of the facts, given his criterion of obligatoriness of both negative elements. However, Van Gelderen's understanding of what constitutes evidence of an ongoing negative cycle seems to be very liberal. For instance, she notes (2008: 230) that certain Berber languages such as Tuareg and Tashelhiyt lack an innovative element that can be analysed as occupying the specifier of NegP (i.e. they have not progressed beyond stage I, as pointed out above in section 1.3.3). However, she states that there is evidence (she gives the example in (26)) that the original negator "soon will be reinforced".

(26) wər ija wæla əndəræn

NEG be.done.PRF.3MSG even little

‘Nothing happened at all.’ (Tamasheq Tuareg (Mali); Heath 2005: 289)

It is hard to see how this statement is justified. The availability of items which serve to emphasize negation, whether these are apparent negative concord items as in (26) (see section 5.2 on Arabic *wala*) or negative polarity items, is ubiquitous in the world’s languages. Various languages can be pointed to (e.g. Slavic languages) which have clearly not undergone Jespersen’s Cycle in several millennia despite having, in common with all other languages, elements which emphasize the polarity of negation, as well as negative concord.

This point aside, one way of shedding light on how common Jespersen’s Cycle is in reality is to take a systematic look at Indo-European languages traditionally spoken in Europe and to count how many have undergone the cycle since Proto-Indo-European.²² It is beyond the scope of this thesis to do this for a large-scale representative sample of the languages of the world, but it makes sense to look at European languages first, since it seems to be the case in general that those who share Van Gelderen’s opinion, do so on the basis of their impression of the situation in Europe.

An indicative summary of this situation is given in Table 1.1, which shows for 47 Indo-European languages of Europe whether or not each one has undergone Jespersen’s Cycle at some point in its history. This is intended to be a near-

²² Jespersen’s Cycle is to be understood for present purposes as the development of at least a stage II construction.

exhaustive sample of Indo-European languages traditionally spoken in Europe, with the following provisos.

There are, of course, many European languages which have undergone Jespersen's Cycle, and many more which have not, and in a number of cases languages falling into one or other of these categories are very closely related (i.e., only split a few centuries ago). Since the parameter under investigation is historical, relatedness of languages in the sample raises a serious risk of bias. Thus we need a principled method of excluding some varieties while including others so as to give a more accurate reflection of the actual situation.

This works as follows. Taking first those closely related varieties which have not undergone Jespersen's Cycle, for each cluster of such varieties I have only included one (arbitrarily selected) representative member in the sample. This is because there might be doubt that the few centuries since the varieties in question split from their common ancestor is a sufficient period in principle for a language to undergo Jespersen's Cycle, and that being the case, each of these clusters should, collectively, only count as one 'no'. Hence, for example, Czech, which has not undergone Jespersen's Cycle, is included in the sample as representative of the Czech-Slovak cluster (and Slovak, which has also not undergone Jespersen's Cycle, is excluded), and this cluster thus receives only one 'no'.

By the same token, in order to avoid artificially inflating the figure for languages that have undergone Jespersen's Cycle, where the cycle occurs in the common ancestor of several daughter languages rather than in the daughter languages themselves, only the common ancestor is counted. Hence the whole of North Germanic only receives one 'yes'. An exception is made, however, for certain closely

related Romance varieties which have nevertheless parted ways with respect to the cycle.

Table 1.1 Jespersen's Cycle in the history of European languages

Language	Has it undergone Jespersen's Cycle?	
	Yes	No
Slavic		
East Slavic		
Russian		1
Ukrainian		1
Belarusian		1
West Slavic		
Polish		1
Czech		1
Sorbian		1
South Slavic		
Serbian		1
Bulgarian		1
Germanic		
East Germanic (Gothic)		1
West Germanic		
High German	1	
Low German	1	
Frisian	1	
English	1	
Dutch	1	
North Germanic (Old Norse)	1	
Italic		
Latin		1 ^a
Romanian		1
Sardinian		1
Italian (Florentine)		1
Napoletano-Calabrese		1
Sicilian		1
French	1	
Franco-Provençal	1	
Friulian		1
Romansch	1	
Trentino		1
Emiliano	1	

Southern Romagnolo		1
Ligurian		1
Lombard	1	
Piedmontese	1	
Venetian		1
Catalan		1
Occitan	1	
Spanish (Castilian)		1
Astur-Leonese		1
Galician-Portuguese		1
Greek		1 ^b
Albanian		1 ^c
Baltic languages		
Latvian		1
Lithuanian		1
Armenian		1
Celtic languages		
Welsh	1	
Cornish		1
Breton	1	
Irish		1
Scottish Gaelic		1
Total	15	32

^a Recall that the univerbation of pre-Latin **ne oenum* into Latin *non* does not fall under the definition of Jespersen's Cycle adopted here.

^b Similarly to Latin, the Greek negative morpheme has undergone a series of univerbations, but at no point has a separate word taken on the function of negator resulting in a bipartite expression of negation (cf. Kiparsky and Condoravdi 2006).

^c Again like Latin and Greek, Albanian has a preverbal negator *nuk*, derived, according to Orel's (2000: 49) reconstruction from univerbation of Proto-Indo-European **ne kʷo* ('not who'). I am not aware of any evidence to suggest that Albanian underwent Jespersen's Cycle as understood here at some point prior to the earliest (15th century) attestations.

(In this table Germanic data come from the references in section 1.2, plus Eythórsson (2002); Romance data come from Poletto (2007) and Parry (2009); and Welsh and Breton data come from Borsley, Tallerman and Willis (2007). Other languages in the table, except those indicated in the footnotes, have self-evidently

retained the stage I construction of Proto-Indo-European, and no one to my knowledge has argued otherwise).

Table 1.1 shows that, of the 47 European languages in the sample, approximately two thirds have not undergone Jespersen's Cycle since they (or their ancestors) split from Proto-Indo-European. Thus, even if we make the assumption that Jespersen's Cycle was an entirely separate development in each of the fifteen 'yes'-languages in this sample, we have the impression of a development that is neither particularly rare nor particularly common.

However, this impression is likely to be misleading. First of all, the fact that every single West Germanic language in the sample has undergone the cycle raises the strong possibility of what Malkiel (1981) calls 'slant': that is, that some specific property of the grammar of Proto-West-Germanic that was inherited by all the daughter languages was responsible for the onset of Jespersen's Cycle in each of them. To the (significant) extent that this sample contains a genetic bias, then, and to the extent that the figures in Table 1.1 misrepresent the history of all the world's languages, the true figure for the pervasiveness of Jespersen's Cycle is likely to be lower than that suggested here, rather than higher.

Secondly, to develop the point made in section 1.2 and Figure 1.1, a clear majority of the 15 'yes'-languages are known to have been in extensive contact with another 'yes'-language during the period of progression from stage I to stage II. (A conservative summary of these contacts: Low German with High German, Dutch with Low German, Frisian with Low German and Dutch, Welsh with English, Breton with French, Occitan with French, Franco-Provençal with French, Piedmontese with French and Franco-Provençal, Lombard with Piedmontese, Emiliano with Lombard,

and Romansch with High German). There must, therefore, be a severe risk that the areal bias of this sample, in addition to the genetic bias, has also contributed to an artificially inflated impression of the pervasiveness of the cycle.

It is sufficient for present purposes to conclude from this, and from Dryer's rather low figures for stage II negation in a properly stratified sample of over a thousand languages, that Jespersen's Cycle should be thought of as neither extremely rare, nor particularly common in the histories of the world's languages.²³

2.3 The case for Arabic-Berber contact

Given this conclusion, it is worth examining in greater detail the historical, dialect-geographical and linguistic data of stage II negation in Arabic and Berber first of all, to see whether they favour a contact-based explanation of this similarity.

The first thing to note is that every Berber variety which today features either an optional or an obligatory stage II construction is spoken in a region where a stage II construction is the default for the local Arabic variety. That is, Kabyle and Shawia in northern Algeria, Central Atlas Tamazight in central Morocco, and Tarifit in northern Morocco. The approximate areas in which these varieties are spoken is indicated in Figure 2.1.

²³ A great deal more research would be required to verify this impression empirically. For Bantu, however, Devos and van der Auwera (2009) give figures for bipartite negation which paint a similar picture to what we have observed in Europe: in a geographically diversified sample of 100 Bantu languages they found 39 with bipartite negation. Many but not all of these are clustered areally.

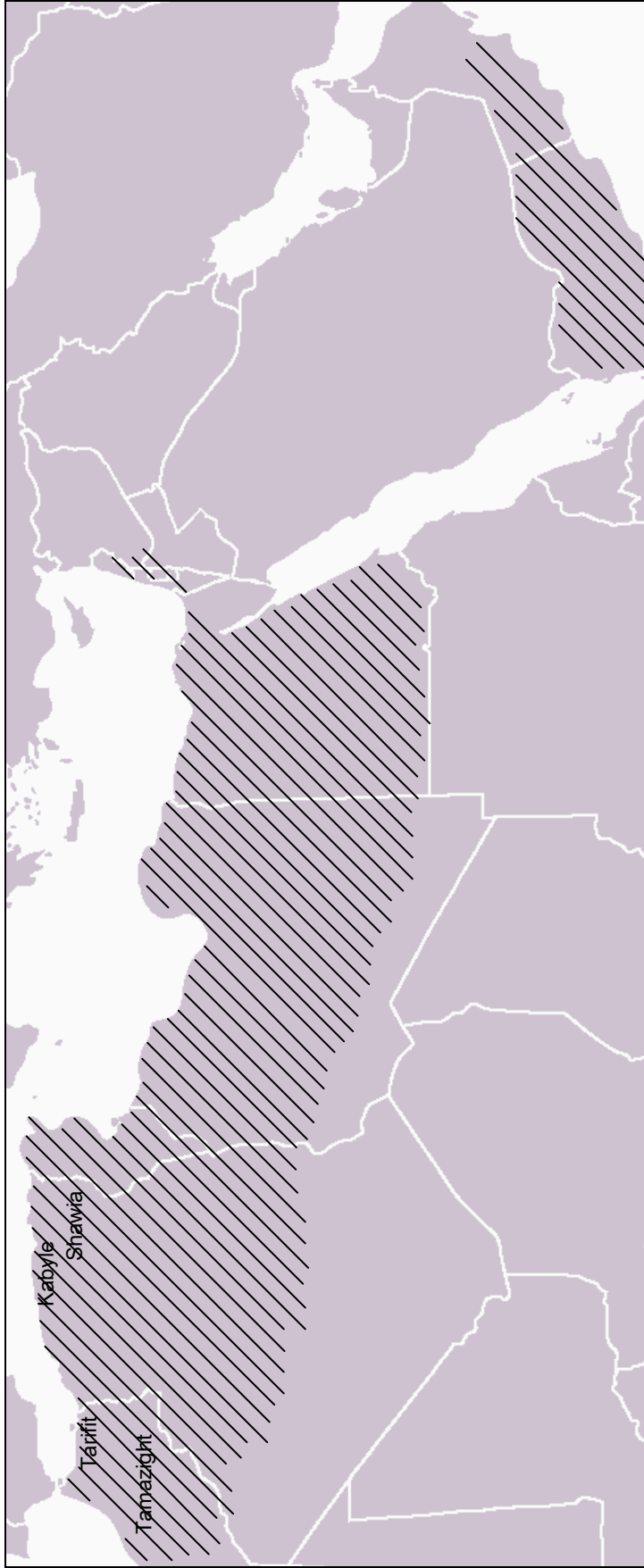


Figure 2.1

Jespersen's Cycle in Arabic and Berber

Shaded areas indicate regions in which a stage II/III Arabic variety is currently spoken. All other regions stage I or non-Arabic-speaking.

Approximate areas in which stage II Berber languages are spoken is indicated by the position of their respective names.

Secondly, the first language of the majority population in these countries is Arabic and there is widespread Arabic-Berber bilingualism among the Berber-speaking minorities. This seems to have been the case for at least some Berber speakers from an early stage in the Arab occupation of the Berber-speaking regions, to judge from the rapid and widespread conversion of the Berbers to Islam and the high degree of social (in particular military) integration between Arabs and Berbers from at least the beginning of the eighth century (Guichard 2000).

Furthermore, we have seen that the etymologies of the postverbal negators in Arabic and Berber are directly parallel, both deriving from words for ‘thing, anything’.

More striking, however, is that there are also close parallels in the grammatical environments in which the postverbal element is obligatorily omitted in all the stage II varieties of north African Arabic and Berber.²⁴

1. With ‘never’ expressions:

(27) **ʔomr-u** ma-ža(*-š)

age-his NEG-come.PRF.3MSG

‘He never came.’

(Tunisian Arabic; Chaâbane 1996: 122)

(28) **ur-jin** diddi (*ša) ɣur-x

NEG-ever come.PRF.3MSG to-us

‘He never came to us.’

(Central Atlas Tamazight; Boumalk 1996: 43)

²⁴ Among the Arabic dialects, the postverbal negator only seems to be obligatorily omitted in these environments from Tunisian westwards. In Eastern Libyan (Owens 1984: 161), Cairene (Woidich 1968: 50) and Palestinian (own data) its omission has become optional.

2. With indefinite pronoun arguments:

- (29) Ma kienu jxewxu(*-x) l-ħadd.
 NEG AUX.PST.3PL provoke.IMPF.3PL to-nobody

‘They weren’t provoking anyone.’ (Maltese; Sutcliffe 1936: 212)²⁵

- (30) ur illi (*ara) yiwən
 NEG there.is anyone

‘There isn’t anyone.’ (Kabyle Berber; Rabhi 1996: 29)

3. Co-ordinate structures:

- (31) ma kla(*-š) ma šrəb(*-š)
 NEG eat.PRF.3MSG NEG drink.PRF.3MS

‘He neither ate nor drank.’ (Morrocan Arabic; Adila 1996: 108)

- (32) u šin (*ša) u swin (*ša)
 NEG eat.PRF.3MPL NEG drink.PRF.3MPL

‘They neither ate nor drank.’ (Tarifit Berber; Lafkioui 1996: 57)

While these similarities are striking, it should be stressed that they are also shared by unrelated stage II languages such as French. On this basis, then, we cannot exclude the possibility that the development of stage II constructions in Arabic and Berber was independent, if, as seems likely, these similarities can be explained

²⁵ I follow here (and throughout) the standard Maltese orthography, the only detail of which that is important for present purposes being that <x> represents /ʃ/.

by general principles. However, there is a further, rather more idiosyncratic environment in which the postverbal negator is obligatorily omitted in these varieties of Arabic and Berber. This is when the negation is in the context of a statement whose truth is strongly guaranteed by the speaker, usually by invoking God as a witness:

(33) **wəllāh** ma-ngūl-ha-lu(*-š)

by God NEG-say.IMPF.1SG-it-to him

‘By God, I won’t tell him it.’ (Moroccan Arabic; Caubet 1996: 86)

(34) **wəlləh** ur t-swiḡ (*ara)

by God NEG it-drink.PRF.1SG

‘By God, I didn’t drink it.’ (Kabyle Berber; Mettouchi 1996: 192)

Taken together, the above considerations strongly suggest that Jespersen’s Cycle in Arabic and Berber were not independent parallel developments. A first step towards the kind of explicit account aimed at here of the role of contact in this shared development is to establish which was the source language and which was the recipient language for the change(s) in question. In the following section I argue that the stage II construction must have been innovated in Arabic and thereafter spread to Berber, and not vice versa.

2.3.1 Did stage II negation first arise in Arabic or Berber?

Before now, others (such as Basset 1952: 37) have also concluded that the stage II construction was innovated in Arabic and then spread to Berber. However, it seems

that this view has principally been based on the misapprehension that the formal similarity of the Berber postverbal marker *ša/-š(i)* to its Arabic counterpart *-š* is a guarantee that Berber has borrowed not only the construction, but also the actual form of its postverbal marker from Arabic, given that the etymology of Arabic *-š* is not in question. As mentioned in section 1.3.3, Brugnatelli (1986) demonstrates that this formal resemblance is in fact coincidental, in that it stems from a regular sound change observable in the majority of Berber dialects with stage II negation, whereby *k > š*, hence *kra > ka > ša > -š* (cf. Chaker 1996: 16).

However, Brugnatelli (1986) goes further than this and in fact turns the traditional view on its head by arguing that stage II negation is an ancient feature of most, if not all Berber languages. His claim is then that especially (if not quite only) those varieties of Berber which underwent the sound change *k > š* were able to maintain their stage II negative construction, supported by the phonetically and syntactically similar Arabic postverbal marker *-š*, whereas the other dialects developed (or perhaps reverted to) stage I constructions. However, this scenario is not at all in accordance with the small amount of extant Old Berber textual evidence, as we will see shortly.

Chaker (1996), on the other hand, is cautious almost to a fault:

“Pour toutes ces raisons, il semble assez difficile de retenir la thèse d’un emprunt direct du berbère à l’arabe ou l’inverse. On doit plutôt envisager une évolution convergente par contact, allant dans le sens la constitution d’une

négation à deux éléments, le second élément étant puisé, dans les deux langues, dans les mêmes classes lexico-sémantiques.”²⁶ (Chaker 1996: 16-17)

It does not follow from the fact that we cannot definitively prove direct borrowing from Arabic on the part of Berber (or the reverse) that this must not have been what in fact happened. Moreover, Chaker provides no positive evidence for ‘convergent evolution’, neither does he explain how or why this might have proceeded.

In fact there is good evidence that stage II negation arose in vernacular Arabic some time between the beginning of the eighth century and the end of the eleventh, most likely in north Africa east of Morocco. This evidence comes from historical facts concerning the movements of Arab and Berber armies and populations in the Mediterranean during this period, together with what we can infer about the nature of vernacular Arabic and Berber at the time, both from what little textual evidence is available, and from contemporary facts about the long-isolated Arabic varieties, Maltese and Cypriot Maronite Arabic.

A few decades after the start of the Arab expansion out of the Arabian peninsula came the first military incursions into Cyprus in around 650 CE. These were quickly followed by fairly large-scale colonization, and in 653 12,000 Arab soldiers and a number of colons from Ba‘albek (in modern Lebanon) settled in Paphos (Kyrris 1985: 187). For the next three centuries Cyprus was ruled jointly by the Arabs and the Byzantine Empire, until the latter took total control around 960. During this time a

²⁶ “For all of these reasons it seems rather difficult to retain the idea of direct borrowing from Arabic into Berber or vice versa. One ought rather to envisage a convergent evolution due to contact, which leads to the formation of bipartite negation, the second element of which is drawn from the same lexico-semantic classes in both languages.”

steady stream of (presumably Arabic-speaking) Maronite Christians from the Levant began to settle in Cyprus, and the last major influx of Maronites came at the end of the Crusades (Dib 1971: 77). Thereafter, the speakers of Cypriot Arabic who remained on the island became largely cut off from the rest of the Arab world, and have remained so until the present day (Borg 2004). In this connection notice that modern Cypriot Maronite Arabic is a stage I language (Borg 2004: 303):

(35) mā ruxt

NEG go.PRF.1SG

‘I didn’t go.’ (Cypriot Maronite Arabic; Tsiapera 1969: 68)

Thus it seems highly likely (and nobody has argued against this to my knowledge) that the vernacular Arabic of pre-1300 Arabia and the Levant was still stage I.²⁷

At the same time, the whole of north Africa, from Egypt to Morocco, was conquered by the end of the seventh century, Egypt having been invaded for the first time in 639. The first campaign against Spain was led from Tangiers in 711 by the Berber commander Tariq bin Ziyad and, again, extensive settlement followed shortly thereafter. It is instructive to note, however, that:

“jamais, en aucune autre province, des éléments autochtones ne furent à ce point associés à l’expansion comme les Berbères le furent pour la conquête de l’Espagne... Du nombre des Berbères qui s’installèrent alors en Espagne, on

²⁷ See section 2.4.3 for the issue of stage II negation in Yemen and Oman.

peut seulement dire qu'il était sans doute très supérieur à celui des Arabes." ²⁸

(Guichard 2000: 30-31)

Importantly for the present argument, the Arabic dialect that this majority Berber population brought with them appears to have used only stage I negative constructions. We can say this with a fair degree of confidence because a large proportion of the few strongly vernacular Arabic-influenced (i.e. 'Middle Arabic') texts that have come down to us were composed in medieval Spain. These were chiefly written by Jews and Christians (presumably because they had less incentive to adhere to the classical standard; cf. Blau 1965) and there is no sign in these texts of a stage II construction (Corriente 1977: 142). This despite the fact that one of the commonest negative constructions in these texts is an evident colloquialism, with a negator *iš* (< Spanish Arabic *aš* 'what' < Classical Arabic *ayy shay?* 'which thing') that is not found in Classical Arabic (or even in any present-day spoken dialects):²⁹

(36) *iš* *tanfaʿ* *al-waṣiyya*

NEG be of value.IMPF.3FSG the-advice

'The advice is useless.'

(Spanish Arabic; Corriente 1977: 145)

²⁸ "At no point, and in no other province were indigenous elements associated with expansion to the extent that the Berbers were for the conquest of Spain... As for the number of Berbers who then settled in Spain, we can only say that it was most probably far higher than the number of Arabs."

²⁹ *iš* is "segregated from the pronunciation [sic] with strong *imāla* of the interrogative *aš*" (Corriente 1977: 145). Note that this is another instance of the change that appears to have created negative *mā* from interrogative *mā* in pre-Classical Arabic.

Hence it seems unlikely that the Arabic dialect spoken in Morocco in the early eighth century and that brought by settlers to Spain in the centuries following could have already progressed to a stage II negative construction by this time. Furthermore, the fact that the majority of settlers were Berbers would seem to cast some doubt on a theory which ascribed stage II constructions in Arabic to substrate influence from Berber, given that there is no evidence of such an influence in Spanish Arabic.

Positive evidence for the development of the stage II construction in Arabic comes from Maltese, which, as we have seen, is today also a stage II language. One might argue that the Maltese stage II construction is the result of an independent innovation, separate from that which resulted in the mainland north African vernacular Arabic stage II construction. However, the near identity of the negative constructions in Maltese and the mainland dialects render such an argument implausible:

1. As already shown, the two elements of the discontinuous morpheme *mā* and *-š(i)* are etymologically and phonologically (almost) identical in all dialects including Maltese.

2. In the mainland dialects, where the sentence includes an auxiliary marker it is to this word that the discontinuous morpheme attaches:

(37)	<i>mā</i>	<i>kān-š</i>	<i>yəḍhak</i>	<i>mʃā-k</i>
	NEG	AUX.PST.3MSG-NEG	laugh.IMPF.3MSG	with-you

‘He wasn’t joking with you.’

(Algerian Arabic; Elhalimi 1996: 143)

The same is true for Maltese:

- (38) Ma **kont-x** smajt l-istoria.
 NEG AUX.PST.1SG-NEG hear.PRF.1SG the-story
 ‘I hadn’t heard the story.’ (Maltese; Borg and Azzopardi-Alexander 1997: 88)

3. Maltese shares the constituent negator *muš* with the mainland varieties:

- (39) **Ħut-ek** **mhux** sejirin għalissa.
 siblings-your NEG go.PTCP.PL for now
 ‘Your siblings are not leaving yet.’
 (Maltese; Borg and Azzopardi-Alexander 1997: 89)

4. It also shares the negative copula formed by cliticization of *mā...-š* negation to the pronouns:

- (40) **M’inti-x** għajjen?
 NEG-you.SG-NEG tired
 ‘Aren’t you tired?’ (Maltese; Sutcliffe 1936: 181)

This suggests that either the Arabic-speaking Muslim settlers in Malta brought with them a stage II dialect, or that these structures spread to the Maltese language from mainland dialects at an early stage in its development. Either scenario is possible, but the former seems more likely. Malta was conquered by the Arabs in

approximately 870. However, according to Al-Ḥimyarī's account (discussed by Brincat 1995), the first significant settlement by Arabs was not until 1048. These settlers most likely came from Tunisia by way of Sicily (Brincat 1995: 22, Cassar 2000: 59). In 1090 Malta was conquered by the Normans and became a fiefdom, under which the Arab community was left relatively intact. The presence of Arab-Sicilian and Tunisian pottery in Malta dating from the eleventh century indicates reasonably strong inter-Arab trade links at this time. However, these links were severed from 1127 onwards when Roger II of Sicily tightened his grip on Malta, and Muslims became increasingly oppressed, the last of them being forced to convert or expelled in 1248 (Cassar 2000: 60). From then on there was very little contact between Malta and other Arabic-speaking countries. Thus the Siculo-Tunisian dialect first spoken in Malta was probably stage II from the outset. If it only became so later due to contact with mainland Tunisian speakers, this would have had to have happened in the relatively short period from 1048 to 1127. Either way, the evidence points to Tunisian having been stage II by the end of the eleventh century at the latest, and possibly several centuries earlier.

Unfortunately, but not unexpectedly, there is no detailed textual record of the dialect of Arabic spoken in this region in this period. The earliest vernacular Maltese text is the *Catinlena* of Pietru Caxaro, a short poem which dates to some time before the death of its author in 1485. This text contains four negative sentences. Two of these involve coordination and so feature a stage I construction as expected. The other two, however, which resemble each other very closely, are simple negative declarative sentences, and both feature a stage II construction. I cite them here in their original orthography and in standard Maltese orthography:

- (41) Mectatilix mihallimin me chitali
Ma ħtat-li-x mgħallmin, 'mma qatagħ-li
 NEG wrong.PRF.3FSG-to.me-NEG builders but cut.PRF.3MSG-to.me
 tafal morchi
 tafal merħi
 clay loose
 'Builders didn't wrong me, rather it was loose clay that gave way.'

- (42) Mectatilix il mihallimin ma kitatili
Ma ħtat-li-x l-ingħallmin, 'mma qatgħet-li
 NEG wrong.PRF.3FSG-to.me-NEG builders but cut.PRF.3FSG-to.me
 li jebel
 l-ġebel
 the-rock
 'The builders didn't wrong me, rather it was the rock that wronged me.'
 (C15th Maltese; Luttrell 1975: 67)

The gender agreement on the verbs in these examples is somewhat puzzling, but the evidence of stage II negation is clear.

We are now in a position to address the question of which of Berber and Arabic was the source language and which was the recipient as far as the stage II negative construction is concerned.

Recall that Brugnatelli's (1986) claim (unsupported by any textual evidence) was that the stage II construction was an ancient feature of Berber which was then lost

(and replaced with a stage I construction) in all varieties except those which were in contact with stage II Arabic dialects. Fortunately, there is a small amount of Berber textual evidence from the twelfth century which is available to us today, and which casts doubt on the tenability of Brugnatelli's claim. This comes in the form of an anonymous Ibāḍī chronicle from the region of Jebel Nefousa in western Libya (where the present-day Berber dialect is stage II; Brugnatelli 1986: 53). The text is principally in Arabic with a number of Berber passages interspersed. It has been published by Lewicki (1934) with a postscript including an edited transcription by André Basset. Among the Berber passages are six instances of sentential negation, all of which are clearly stage I; for example:

(43) ur tšnt trā
 NEG be.erased.IMPF.3FPL scriptures
 'The scriptures are not erased.'

(44) ur tšn ĩtrān
 NEG be.erased.IMPF.3MPL stars
 'The stars are not erased.'

(C12th Jebel Nefousa Berber; Lewicki 1934: 300)

Of course, a lack of evidence is not evidence of a lack, and it cannot be excluded that this variety of Berber had the option of a stage II construction in contexts other than the ones found in this text. However, this is sufficient evidence to lead us to reject Brugnatelli's (1986) strong claim that stage II negation was the norm in Berber prior to the Arab conquest of north Africa. The earliest evidence of a stage II

construction in Berber that I am aware of comes from a Kabyle word and phrase list published by Shaw (1738: 52). Fortunately, the one example he gives of negation is with a stative verb, demonstrating that the postverbal *kra* cannot be an argument meaning '(any)thing':³⁰

(45) neck urfedaag ikra
 nekk **ur** fadagh **kra**
 I **NEG** thirst.IMPF.1SG **NEG**
 'I am not thirsty.' (early-C18th Kabyle Berber; Shaw 1738: 52)

To sum up then, the fact that stage II negation is absent from the Cypriot and Medieval Spanish dialects, but present in Maltese, indicates that the innovation took place in Arabic not earlier than the beginning of the eighth century and certainly not later than the end of the eleventh. We also have a twelfth century text in a Berber dialect which appears to be stage I and which was spoken in the same general region (central north Africa) as an Arabic dialect (Tunisian) which we have seen must have been stage II by that time. This, taken together with the fact that the present-day varieties of Berber that are stage II are largely restricted to those regions where the local Arabic dialect is also stage II (see section 2.2), as well as the marked similarities in the stage II constructions in both languages, all points to the conclusion that the stage II construction was an innovation that began in Arabic and spread through language contact to Berber. On this basis, a natural question to ask next is, in which dialect(s) of Arabic did this innovation take place, and what was the cause?

³⁰ I am indebted to Lameen Souag for pointing out this example to me.

2.4 The case for Coptic-Arabic contact

Let us consider again the data on negation in the various present-day colloquial Arabic varieties, given in (6)-(14), and the distribution of the stage II construction among these varieties, illustrated in Figure 1.2. In considering data such as these, a number of different levels of explanation of the changes in question are possible. At the first level, we can make suggestions as to the likely etymology of the innovative element and, where possible, draw parallels with languages whose histories are better documented where a comparable element has undergone a comparable development. This is the level that almost all mention in the literature of the development of negation in Arabic has been content to stop at: it seems clear that one of the many outcomes of the grammaticalization of Classical Arabic *šay?* ‘thing’ is the negative enclitic *-š*. Concerning some of the changes addressed in this thesis (especially in chapter 6) it will not be possible to go much further than this level of explanation. It is certainly possible in principle though.

The next level involves outlining the necessary conditions for the development to take place – in this instance the mechanisms of syntactic reanalysis and phonological attrition, and examples of the type of bridging context that would have made the reanalysis possible. I gave an explanation of the Arabic data along these lines in a (2007) article, in which I argued that the reanalysis took place when speakers produced *šay?/š?* as the argument of a negated verb in discourse contexts where the acquirer could have reasonably expected this argument to be null. We revisit this hypothesis in section 4.1.

The important point for present purposes is that there is a still deeper level of explanation possible. In addition to outlining the necessary conditions for a reanalysis to take place, we can attempt to specify the sufficient conditions. That is, we can attempt to answer Weinreich, Herzog and Labov's (1968) 'actuation problem':

"For even when the course of a language change has been fully described and its ability explained, the question always remains as to why the change was not actuated sooner, or why it was not simultaneously actuated wherever identical functional properties prevailed. The unsolved actuation riddle is the price paid by any facile and individualistic explanation of language change. It creates the opposite problem – of explaining why language fails to change." (Weinreich et al. 1968: 111-112)

In the case of the Arabic Jespersen's Cycle we can ask: what extra factor(s) obtained in north Africa and the southern Arabian Peninsula such that this change actually occurred and spread, but were absent outside these areas such that there it did not? Sometimes when this kind of question arises we will be able to point to prior changes, unique to the grammars of the innovating varieties, that led to these varieties alone undergoing the change under investigation. This cannot be done for Jespersen's Cycle in Arabic owing to the lack of significant attestation of earlier spoken varieties. In principle, it cannot be ruled out that something along these lines is at work in this case. However, the argument we will explore here is that the uneven distribution of Jespersen's Cycle in the Arabic dialects is best explained as resulting

from contact between Coptic and Arabic as spoken in Egypt, and between Modern South Arabian and Arabic as spoken in Yemen and Oman.

Let us start by considering a number of apparent obstacles for such an account. Recall from examples (18) and (19) that a bipartite construction is a central feature of the Coptic negation system. In and of itself, however, the fact that Coptic and certain varieties of Arabic both have stage II constructions clearly tells us nothing about the role of the former in the development of the latter, despite the fact that we know that there was intensive contact between Coptic and Arabic in Egypt. (See Bishai 1962 for some further examples of Coptic grammatical influence on Egyptian Arabic).

Moreover, the resemblance could be seen as merely superficial: note that the placement of the Coptic postverbal negator *an* is rather different to that of its Arabic counterpart *-š*. In Arabic verbal sentences *-š* always cliticizes to the end of the verb, as in (7)-(9), with only direct and indirect pronominal object clitics potentially intervening. This word order is attested in Coptic, as in (18)b, but the more usual order is illustrated in (18)a, with *an* occupying a clause-final position.

Furthermore, from what we know of the etymology of *an*, it is quite different to that of Arabic *-š*: *an* is generally agreed to be descended from an item *iwn3*, which first appears in Late Egyptian texts, where it functions as a strong negative polarity adverb 'at all' (Gardiner 1904). Its earlier history is unknown (but see Groll 1970 and Davis 1973 for some suggestions, which are, however, rejected by Meltzer 1990). As such, Arabic *-š* (< *šī* < *šay?* '(any)thing') cannot be a calque on Coptic *an*.

These difficulties are addressed in section 2.2.4. First, however, we must address what is potentially the most serious obstacle of all – the possibility that Coptic may no longer have had a stage II construction in the period of contact with Arabic. This

possibility is suggested by the fact that from an early period we find stage III negative constructions in Coptic, illustrated in (46) (= (20)) appearing alongside the more conservative stage II construction.

(46)	anon	an	pe-nt-a-n-mere	p-noute
	we	NEG	COP-REL-PRF-1PL-love	the-God
	'It is not we who loved God...'			(I John 4:10)

That is, the innovative Coptic postverbal negator *an* has become able to express negation alone, without the obligatory presence of the original preverbal negator *ən*, which has become optional, at least in some contexts.

On this issue, Kickasola (1975: 272) claims, without referring to any specific time period, that the first negative element in the stage II construction, *ən*, can be omitted wherever it can occur. This appears already to be true several centuries before the arrival of the Arabs in Egypt, to judge from third- or fourth-century examples such as (46),³¹ where the focus of negation – the clefted subject *anon* ‘we’ – is not preceded by *ən*. As such, it is crucial to the present argument to know whether the stage III construction had increased significantly in frequency at the expense of the older stage II construction by the time the Arabs arrived in the seventh century and the Copts started learning Arabic in large numbers in the following centuries.

If the stage III construction had become unmarked in the relevant contexts by the time large numbers of Copts began learning Arabic as a second language in the

³¹ On the dating of the Sahidic New Testament, see the entry ‘Coptic versions of the Bible’ in the Catholic Encyclopedia: <http://www.newadvent.org/cathen/16078c.htm>.

centuries following the initial invasion in 639, it is hard to see how or why they should have imposed a stage II construction on their second-language Arabic. Indeed, studies of the second-language acquisition of negation such as Stauble (1984) and Meisel (1997) show that native speakers of VNeg languages such as Japanese do not differ from native speakers of NegV languages such as Spanish with respect to their placement of negation in their second-language English, which is preverbal in all cases. That is, speakers of stage III languages do not appear generally to impose a stage III construction on a stage I second language.³² By the same token, we would not expect speakers of a stage III language to impose a stage II construction on a stage I second language either. Thus it is crucial to ascertain whether Coptic in its late stages was in fact solidly stage III in the relevant negative constructions, that is, whether the original preverbal marker had become obsolete. The corpus study presented in the following section was designed to ascertain whether this was in fact the case.

³² There may be exceptions to this generalization, particularly in cases of radical underexposure to the lexifier language in pidgin and creole formation. Jonathan Owens (personal communication) points out that the Ugandan variety of East Africa Creole Arabic, for example, has a stage III construction with clause-final *ma*, probably as a result of substrate influence. Of course, to the extent that native speakers of stage III languages *can* impose stage II or III constructions on stage I L2s, doubt that Copts could have been partly responsible for Jespersen's Cycle in Arabic is reduced still further.

*2.4.1 Corpus study of negation in late Coptic*³³

For the purposes of this study, a corpus of Coptic texts from the ninth and tenth centuries was compiled. This time period was chosen in order to reflect the language spoken in Egypt at the beginning of the Arab dominance in the area. It should be stressed from the outset, however, that in trying to gain an insight into the state of the spoken language of this period, we face the familiar problem that the texts used are likely to represent a more conservative register than one would have found in everyday speech. That said, native speakers of Coptic would have been learning Arabic as a second language for several centuries by the time these texts came to be written down. So it is perhaps no bad thing, for current purposes, if they represent to an extent the speech of earlier generations.

The texts are all religious in nature and written in Sahidic, a dialect spoken originally in the region of El-Ashmounein in the central part of the modern state of Egypt. This dialect was the main literary variety from the third to the ninth centuries and survived until the fourteenth when it was replaced by Bohairic, the standard language of the modern Coptic Church.

The texts used are as follows. First, the Martyrdom of St Coluthus (C), Paese and Thecla (PT) and Shenoufe and His Brethren (SB). The manuscripts of these texts contain colophons explicitly dating them to the mid-ninth century (Reymond and Barns 1973: 20). Additionally, three translations from Greek of Basil of Caesarea's sermons were used, namely, the Encomium on St. Mercurius (M.B) (explicitly dated

³³ Similarly to section 1.3.2, the data presented in this section were obtained in collaboration with Elliott Lash. The construction of the corpus, the collection and interpretation of negative sentences from that corpus, and the production of the figures relating to these sentences were all done jointly. As with the data in section 1.3.2, the data in this section are to be published in Lucas and Lash (forthcoming).

842), The First Homily on St. Michael Archangel (MA.B) and the Second Homily on St. Michael Archangel (2MA.B), both of which can be confidently dated to the ninth or early tenth centuries (Depuydt 1991: viii). Finally, three translations also from Greek of sermons by John Chrysostom were also included. These were the Encomium on the Bodiless Creatures (BC.J) (explicitly dated 892-3), The Homily on St. Michael Archangel (MA.J) (9th / early 10th century) and the Homily on the Resurrection and the Apostles (RA.J) (explicitly dated 855).

From each of these texts, all of the negative sentences were excerpted and categorized according to negation type. The main focus of the study, however, was to track the presence or absence specifically of the first element in the bipartite *ən...an* construction. If the stage II construction was still alive and well in the first-language Coptic of second-language learners of Arabic, then this should result in the proportion of occurrences of *an* with a preceding *ən* in the corpus being greater than that of *an* with no preceding *ən* (i.e., a stage III construction).

This was indeed the finding for all texts in the corpus; no text had more occurrences of *an* without *ən* than with it. The following table gives the relative percentages of each of these two variants in the corpus.

Table 2.1 Percentages of stage II and stage III negative constructions in the Coptic corpus

	C	PT	SB	M.B	MA.B	2MA.B	BC.J	MA.J	RA.J	Total
% <i>ən...an</i>	100	85	97	67	54	80	55	100	56	83
% <i>an</i>	0	15	3	33	46	20	45	0	44	17
Total tokens #	19	40	34	6	13	10	11	11	9	153

The stage II construction is thus much more frequent than the stage III construction even at this stage. What, then, explains the persistent presence of a small proportion of the latter in almost all texts?

It was mentioned above that Kickasola (1975: 272) claims that the stage III construction (*an* alone) may be found wherever the stage II construction (*ən...an*) would have been possible. He also notes, however, that the distribution of these two alternatives is not entirely random. Although he does not provide statistics, he claims that *ən* is more often omitted before nominal subjects than pronominal subjects. Where it is omitted before pronominal subjects, he notes that this occurs most often before *se* 'they', *ti* 'I', and *tetən* 'you' (plural). This is arguably a phonological phenomenon as all of these begin with coronal consonants, to which the coronal in *ən* might be expected to assimilate.³⁴ Kickasola also notes that *ən* is omitted almost without fail before the imperfect-tense auxiliary *ne(re)* and is apparently inadmissible before the second perfect tense auxiliary, *ənta* (see (49) below). These again appear to be phonologically motivated omissions.

In the present corpus, it turns out that every instance of a stage III construction where a stage II construction would have been expected is similarly attributable to phonological considerations.

First of all, the majority of the few cases of *ən*-dropping found in the corpus are before a coronal nasal /n/. This is true whether the /n/ is a prefix on, or an integral part of, the following word. Thus examples (47) and (48) below have initial /n/

³⁴ The full story may be more complex, however, because Kickasola does not mention the frequency of its occurrence (or lack thereof) before *tən* 'we'.

immediately following the position where one would have expected *ən* to occur: before the initial /n/ of the subject *nekbasanos* in (47) and before the initial /n/ of the auxiliary *ne-* in (48). Of the 17 instances of *(ən)...an* negation in the context of a following /n/ found in this corpus, 15 (88%) featured *ən*-dropping. In the following examples, *(ən)* indicates the position where this morpheme would be expected, were it present.

(47) (ən) nek-basanos na-eš-cəmcom ero-i an
 (NEG) your-tortures FUT-able-find.strength to-me NEG
 ‘Your tortures will not be able to have power over me.’ (PT 66Ri)

(48) (ən) ne-u-na-paht-əf an ejō pe
 (NEG) IMPF-3PL-FUT-pour-it NEG over.2FSG PRT
 ‘It would not be poured over you.’ (PT 70Vi)

Another common context for *ən*-dropping was before syllables beginning with the homophonous sound /ən/. The two main contexts for syllable-initial /ən/ are with the negative of the second perfect, whose auxiliary is *ənta*, and before second and third person independent personal pronouns, which all begin with the sound /ənt/. Examples (49) and (50) show these two contexts. The negator *ən* was dropped in all six of the cases in the corpus where the following syllable began with /ən/.

(49) (ən) ənta-i a’a-u an hən oumənt-magos
 (NEG) PRF-1SG do-them NEG by magic
 ‘It’s not by magic that I did them.’ (SB 119Rii)

(50) (əŋ) əntetno an

(NEG) you NEG

‘It’s not you.’

(M.B 1)

The third context for *ən*-dropping was following the morpheme *e*, which is the circumstantial complementizer used to introduce clauses of attendant circumstance. In this case one would have expected the pre-verbal negator, were it present, to appear as *n* directly before the subject. Example (51) shows an attestation of this type from the corpus. There were ten instances of *(ən)...an* negation in this context in the corpus; *ən* was dropped in three of these and retained in seven.

(51) ouəŋoun e (n) k-ər ša an

certainly COMP (NEG) 2MSG-do celebration NEG

‘While you are certainly not celebrating...’

(MA.B 21.3)

Finally, there are two examples in the corpus of omission of *ən* before a subject pronoun beginning with a non-nasal coronal, as against 59 cases of *ən*-retention in the same context. A case of *ən*-dropping is shown in example (52).

(52) (əŋ) ti-ouōš an

(NEG) 1SG-desire NEG

‘I do not want...’

(MA.B 6.3)

Table 2.2 summarizes the environments in which *ən* is retained and where it is dropped. The picture that emerges is that *ən*-dropping appears to be primarily a phonological process, at least in origin. *ən* only seems to be dropped in the immediate environment of sounds with which it shares place and/or manner features.³⁵ Furthermore, the only environment that appears to be actually conducive to *ən*-dropping is before nasals (including nasals immediately preceded by schwa). All other environments favour *ən*-retention.

Table 2.2 Environments for *ən*-retention and dropping

	Pre-Nasal	Pre- <i>ən</i> [tʃ], etc.	Between <i>e</i> and Verb	Pre-Coronal	Other	Total
# <i>ən...ən</i>	2	0	7	59	59	127
# <i>...ən</i>	15	6	3	2	0	26

Thus the Coptic stage II construction with *ən...ən* clearly predominates with respect to its stage III counterpart (which may have no syntactic reality, being merely phonologically conditioned); and this even in the ninth/tenth century, several centuries after the Arab conquest of Egypt and the time when native speakers of Coptic would have begun learning Arabic as a second language.

³⁵ The tokens of *ən*-dropping following the complementizer *e* perhaps have a non-phonological cause, which cannot be explored further here. Suffice it to say that insofar as *ən*-dropping can occur independently of phonological considerations in these texts, this is a distinctly marginal possibility.

2.4.2 Further obstacles for the case for Coptic-Arabic contact

In section 2.3 I showed that the scenario whereby stage II negation spread to Berber through contact with Arabic is plausible from a linguistic, distributional and historical point of view, and is not falsified by data of any of these three types. For the argument for the Coptic origin of Arabic stage II negation to be taken seriously, it needs to satisfy the same criteria. Initially, however, there might be doubt on all three counts.

To start with the structural differences between the Coptic and Arabic stage II constructions, we have already noted that the position of the innovative element in the Arabic stage II construction is fixed in verbal sentences – it is enclitic on the verb, attaching directly to it or to any direct or indirect object pronominal clitics which may intervene:

- (53) mziyya mā gāl-ha-lū-š rājjil
 luckily NEG say.PRF.3MSG-it-to.him-NEG man
 ‘Luckily it wasn’t a man that told him it.’
 (Algerian Arabic; Elhalimi 1996: 138)

In Coptic, however, the position of the innovative element *an* is much freer (though always appearing after the element it negates), and, unlike Arabic, appears generally to follow any nominal direct objects. These appear directly after the verb (which appears in a special pre-nominal form) or directly after a case-marking preposition. If

the object is a pronoun, it may be suffixed to the verb or it may be suffixed to a case-marking preposition. This is shown in (54)a,b.³⁶

- (54) (a) $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$
 NEG 3MSG-desire the-death for NEG
 ‘For he does not desire death.’ (MA.J 20.6)
- (b) $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$
 NEG 2MSG-FUT-find ACC thorns NEG
 ‘You will not find thorns.’ (2MA.B 5.1)

In contrast to the general trend, however, in the largest of the texts in the corpus assembled for this study, Paese and Thecla (PT), we find a sizeable proportion of attestations (29%) in which *an* precedes the verbal complement (whether an NP or a PP):³⁷

³⁶ Examples (54) and (55), as well as the data in footnote 37, were also selected in collaboration with Elliott Lash and will be published in Lucas and Lash (forthcoming).

³⁷ *an* also shows variable order with regard to prepositional adjuncts, which may appear to the left or the right of the post-verbal element:

- i. $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$
 NEG 1SG-FUT-able-keep silence-my to-it NEG
 ‘I will not be able to keep silent about it.’ (M.B 8.2)
- ii. $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$ $\text{\textcircled{a}}$
 NEG 1SG-FUT-keep silence-my NEG to-it
 ‘I will not keep silent about it.’ (M.B 13.2)

- (55) (a) $\text{\textasciix} \text{ ti-sooun} \quad \mathbf{an} \quad \text{\textasciix} \text{ pe-hoou}$
 NEG 1SG-know NEG ACC the-day
 ‘I do not know the day.’ (PT 82Vi)
- (b) $\text{\textasciix} \text{ ti-mpša} \quad \mathbf{an} \quad \text{\textasciix} \text{ pei-noc}$
 NEG 1SG-be.worthy NEG of this-honour
 ‘I am not worthy of this honour.’ (PT 70Vi)

These data are suggestive of a move on the part of the ninth-/tenth-century variety of Coptic studied in this corpus towards a higher position in the clause for *an*. This process, whereby a newly grammaticalized negator descended from an adverbial element (as seems to be the case for Coptic *an*) starts off late in the clause but over time is more and more frequently found close to the verb, is familiar from the histories of European languages such as French (Hirschbühler and Labelle 1994) and Welsh (Willis to appear).³⁸ If we are to maintain that the reanalysis of *šay?* as part of a stage II construction in Arabic was in fact triggered by its somewhat different Coptic counterpart, a number of alternative possibilities, illustrated in Figure 2.2, thus suggest themselves.

First (scenario 1), it could be that native speakers of Coptic already felt in the period in which they started learning Arabic as a second language that the default word order in Coptic negative sentences was for *an* to precede NP and PP

³⁸ This gradual migration of a clause-late negator to immediately postverbal position need not be inevitable, however, particularly where the negator in question is of the ‘resumptive’ kind found, for example, in Brazilian Portuguese, Hausa, and Modern South Arabian (see section 2.4.3).

complements of the verb. Under such a scenario they could then be said to have faithfully replicated this pattern in their second-language Arabic.

Second (scenario 2), we could suppose that native speakers of Coptic at this time in fact still felt that the default order was for verbal complements to precede *an*, and that, again, they faithfully replicated this pattern in their second-language Arabic. However, in the centuries following this transfer Arabic followed French, Welsh and Coptic in increasingly raising its postverbal negator beyond verbal complements until adjacency to the verb became obligatory as it is now.

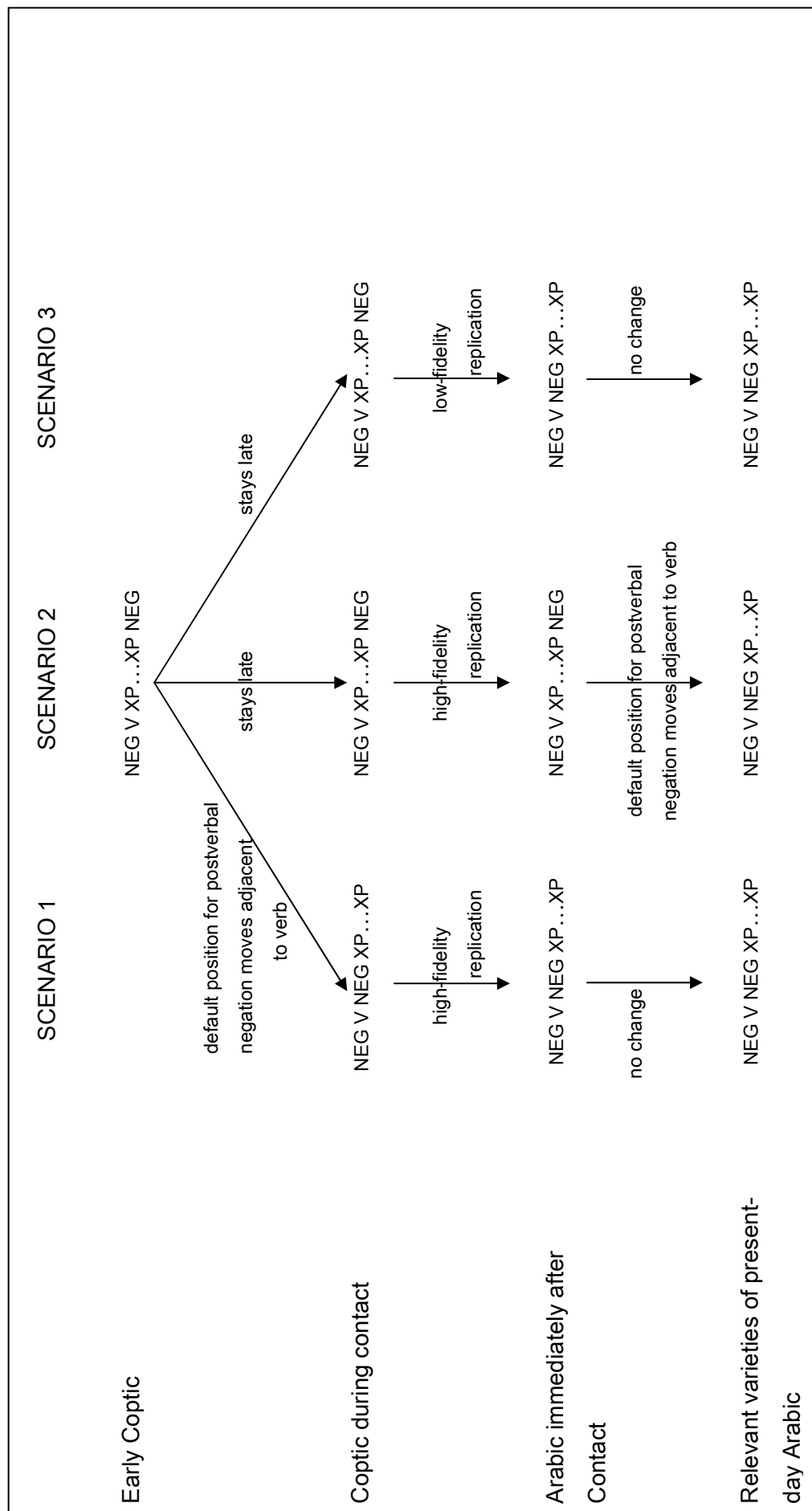


Figure 2.2

Three scenarios for the transfer of the stage II construction from Coptic to Arabic

Finally (scenario 3), perhaps most plausibly, it could be that native Coptic speakers did not in fact faithfully replicate the principal Coptic pattern (with a clause-late negator) in their second-language Arabic (with a verb-adjacent negator). What they instead did was to interpret the Arabic they were exposed to in such a way that they found evidence in it for a stage II negative construction familiar to them from their native Coptic, with the difference that in Arabic (or the version of it that they learnt at least) there was no optionality in the positioning of the postverbal element: it had to be verb adjacent. Under this scenario (which is developed in section 4.1) contact plays an important role in the development of the Arabic stage II construction, but this does not entail the perfect replication of the Coptic structure in Arabic.

Linguists in the tradition exemplified by the quotations from Sapir (1921) and Lass (1997) given at the start of section 2.1 are liable to object at this point that the above scenarios may well be plausible, but, given that the stage II constructions in Coptic and Arabic are not entirely congruent, a more economical hypothesis would state that the one had no influence on the other. Note, however, that there is a high price to be paid for excessive skepticism here. If we deny that Coptic learners of Arabic could have reanalysed *šay?* as a negator, then we are forced into the paradoxical position that first language acquirers, by contrast, were capable of this, since this reanalysis must have taken place at some stage. This is despite the fact that first language acquirers of Arabic would have lacked the specific motivation for positing a bipartite construction that native speakers of Coptic would have had, and that children evidently tend to converge on a grammar of their first language that is a great deal closer to that of other native speakers of that language than do adult second

language learners. Moreover, we face the difficulty of explaining why first language acquirers of Arabic outside the areas in which we find bipartite negation today were not similarly capable of making this reanalysis.

That is, we are faced with a rather clear instance of Weinreich et al.'s (1968) actuation problem. If we wish to offer an account of the development of stage II negation in Arabic that is not "facile and individualistic" (as Weinreich et al. 1968: 112 put it), then we must consider why it is that some Arabic dialects have participated in this innovation while others have not.

If we consult again Figure 1.2, we see that stage II negation is found in a contiguous region across the more northerly part of north Africa, spilling out a little way into the Levant, and then, somewhat incongruously, in a second region in the southernmost part of the Arabian Peninsula (parts of modern Yemen and Oman). Let us first address why it is that stage II negation is not found throughout the Arabic-speaking region, and then why it is found precisely in those regions just mentioned.

As we have seen, the Arabic postverbal negator *-š* is ultimately derived from the Classical Arabic word *šay?* 'thing'. Several features of the syntax of Classical Arabic make it look like a rather fertile breeding ground for Jespersen's Cycle. Firstly, Arabic has never had an indefinite article. As such, a non-referential or quantificational interpretation of a noun of minimal quantity in Classical Arabic, such as *šay?*, is available synchronically in negative polarity contexts (56)a, while in affirmative contexts apparently the same item retains its referentiality (56)b:

- (56) (a) xalaqtu-ka min.qablu wa-lam taku **šayʔan**
 create.PRF.1SG-you before and-NEG be.JUSS.2MSG anything.ACC
 ‘I created you before, when you were not anything (/nothing).’
 (Qur’an 19: 9)
- (b) laqad jiʔta **šayʔan** nukran
 PRT bring.PRF.2MSG thing.ACC horrible.ACC
 ‘You have done a horrible thing.’
 (Qur’an 18: 74)

Secondly, already in the Classical period we find examples of *šayʔ* apparently functioning as an adverb, rather than an argument:

- (57) lā yaḍurru-kum kaydu-hum **šayʔan**
 NEG harm.IMPF.3MSG-you cunning.NOM-their at.all
 ‘Their cunning will not harm you at all.’ (Qur’an 3: 120)

It would appear, then, that by this stage *šayʔ* had already split into two (homophonous) items, the nominal use illustrated in (56) and the adverbial use in (57) (see section 5.3.5 for further discussion on this point). This split would have been facilitated by the fact that the suffix marking indefinite accusative case in nouns is identical to the suffix that marks adverbs: *-an*.^{39, 40}

³⁹ It is a moot point as to whether the case and definiteness inflections of Classical Arabic were also a part of any of the Classical-era spoken varieties from which the modern spoken varieties are descended (see Owens 2006 for the arguments against). In the case of *šayʔ* in particular, however, the evidence of Maltese, where, as we will see in chapter 5, the word for

Moreover, it turns out that the form *šayʔan*, whether functioning as an argument or an adverb, is found predominantly in the context of negation already in Classical Arabic. In the Qurʾan, for example, which consists of approximately 80,000 words, *šayʔan* occurs 77 times. Of these, fully 63 (81.8%) occur in the scope of negation (cf. Talmon 1999). As such, already at this stage, *šayʔ(an)* is an example of what Hoeksema (1994) calls a semi-NPI: an item that predominantly, but not exclusively, occurs in negative polarity contexts. Given this strong association with negation already in Classical Arabic, it would not then come as a surprise to find either the indefinite pronoun or the negative polarity adverb *šayʔ* grammaticalizing as a new postverbal negator, and this does appear to be what has happened in the stage II varieties spoken in the regions indicated in Figure 1.2.

It is clear, however, that these features of Classical Arabic cannot represent sufficient conditions for the future development of a stage II negative construction, because not all of the modern dialects descended from Classical Arabic (or, perhaps more accurately, the Arabic spoken in the classical period; Owens 2006) have developed such a construction. A range of varieties have developed other functional items from Classical Arabic *šayʔ*, but again none of these can be seen as sufficient to trigger the subsequent development of a stage II negative construction, because they

‘nothing’ is *xejn*, suggests that this item at least did occur with the *-an* ending in early colloquial Arabic varieties. For convenience, however, I will continue to refer to the ancestor of postverbal *-š* simply as *šayʔ*.

⁴⁰ It is also unclear whether the homophony between the indefinite accusative and adverbial suffixes is coincidental or the result of the latter being derived from the former. Arabic orthography indicates a derivational link, but further research is required to establish whether this is accurate or merely a folk etymology.

are found both in some dialects which have undergone Jespersen's Cycle and in some which have not.

For example, in several dialects we find a clause-final element *šī* (< *šay?*) which marks interrogation (not negation). It is found both in Syrian Arabic (Cowell 1964) and Cypriot Maronite Arabic (Tsiapera 1969), which have not undergone Jespersen's Cycle, and, among those which have, in Libyan Arabic at least:

- (58) *šift* *aḥmad* *amis* *šī*
 see.PRF.2MSG Ahmad yesterday Q
 'Did you see Ahmad yesterday?' (Libyan Arabic; Owens 1984: 102)

Similarly, *šī* (< *šay?*) is found with the function of an adnominal quantifier 'some X' in both Moroccan Arabic, which has undergone Jespersen's Cycle, and Syrian Arabic, which has not.⁴¹ (For more on the polygrammaticalization of *šay?* in the

⁴¹ Davies (1981) hypothesizes that precisely this adnominal quantifier use of *šī* was reanalysed as a negator (and interrogative marker; see chapter 5) in Egyptian Arabic. However, there are several problems with this proposal over and above the (actuation) problem of why Syrian Arabic, which has the same quantifier use of *šī*, has not also developed a stage II negative construction. First, the change Davies proposes – [Neg V [Q NP]] > [[Neg V Neg] NP] – appears not to be attested in the history of any other language, though the reverse is relatively common (cf. Breitbarth, Lucas and Willis forthcoming). Second, quantifier *šī* is not actually attested in present-day Egyptian Arabic, and there is no strong evidence to suggest that it ever was. Davies (1981: 274) points to its supposed survival in the clearly related quantifier *išī*, but the status of this item in Egyptian Arabic is marginal at best, being restricted to contexts where a number of items are listed (Hinds and Badawi 1986: s.v. *šy*). Finally, Davies (1981: 274-5) states that all the supposed examples of quantifier *šī* in the 17th century Egyptian Arabic material he analyses are found in negative and interrogative contexts, and, what's more, they can all in fact also be interpreted as negative or interrogative verbal enclitics. Given these considerations, it seems unlikely that Davies' hypothesis is on the right track.

Arabic dialects see Opler 1975; on polygrammaticalization in general see Robert (2005)).

For a contact-based account of the development of Jespersen's Cycle in Arabic to be discounted, there needs to be a coherent internal account which could replace it. The onus is on an internal account, which claims that various features of Classical Arabic were sufficient to cause this development, to explain why they were only sufficient for a subset of the dialects descended from Classical Arabic (or something like it). In the absence of such an explanation, the contact-based explanation suggested here seems to be more warranted in view of the facts available. Indeed, this contact-based explanation is given stronger support once we turn to consider why it is that the dialects of Arabic which have undergone Jespersen's Cycle are distributed the way they are.

Their distribution in the region stretching from Morocco to Palestine and neighbouring areas of countries surrounding Palestine is unproblematic if we assume an Egyptian origin, followed by spread via dialect contact: Jespersen's Cycle has spread westwards from Egypt throughout coastal north Africa, following the prevailing flow of migration of Arabic-speaking peoples in this region and consistent with the sphere of influence of Cairo as the capital of Fatimid empire in the tenth to twelfth centuries. Recall from section 2.3.1 that the evidence from Spanish Arabic suggests a relatively late date for Jespersen's Cycle in Moroccan Arabic, which is entirely in keeping with a scenario that sees the stage II construction spreading gradually westwards from Egypt. Similarly, it makes sense that this construction should have

spread only a little way eastwards into the Levant as this is against the prevailing flow of migration (Taha 1989).⁴²

In addition, the timescale for the innovation of stage II negation in Arabic proposed in section 2.3.1 – some time between the beginning of the eighth and the end of the eleventh centuries – is consistent with that innovation having been driven by native speakers of Coptic gradually shifting to Arabic in the period between the arrival of the Arabs in the mid-seventh century and the eventual extinction of Coptic as a spoken language by the sixteenth century at the latest.

A point that should be addressed here, however, is whether this timescale is not potentially falsified by what little textual evidence exists of pre-modern colloquial Egyptian Arabic. The majority of this comes from the Cairo Genizah, which is a body of over 200,000 manuscripts dating from between the ninth and the nineteenth centuries, the majority of which are in the Taylor-Schechter collection held by the University Library in Cambridge. Many of these are in colloquial-influenced ‘Middle Arabic’ (written in Hebrew script) and are therefore potential of relevance to the present argument. Wagner (2007) has analysed a large corpus of eleventh-to-nineteenth-century Judaeo-Arabic letters from this collection and includes a detailed discussion of negation.

In considering this corpus, however, it should be stressed that what we are dealing with here is not an accurate representation of colloquial Arabic in this period. As Wagner (2007: 4) points out, the language of the letters in this corpus, despite

⁴² Moreover, the general trend for Jespersen’s Cycle to be less advanced in Berber varieties the further west and south one travels from Libya and Tunisia (cf. the discussion around examples (21)-(25)) is consistent with a gradual spread westwards and southwards of the cycle in the local contact varieties of Arabic, as suggested above.

having many vernacular features “also exhibits some features that are distinctive of Classical Arabic. Apart from those, they also exhibit some phenomena that are neither vernacular nor Classical Arabic features, but have to be qualified as part of substandard letter writing.” Unfortunately, in the vast majority of this corpus negation is one of the domains that is strongly influenced by Classical Arabic. We can say this with confidence, because the commonest negator in the corpus as a whole is *lam* (Wagner 2007: 174; cf. section 1.3.1), a form which it seems doubtful was ever a part of post-Islamic vernacular Arabic. Moreover, the use of *lam* increases in frequency in the later texts in the corpus, such that it is virtually the sole negator in eighteenth- and nineteenth-century texts, albeit used with the perfect to negate the past and the imperfect to negate the present and future, a hypercorrection with respect to correct Classical usage, in which *lam* can only be used with the apocope (or ‘jussive’) form of the imperfect to express negation of the past. Thus, it has little bearing on our understanding of negation in pre-modern Egyptian colloquial Arabic that there is only one clear case of bipartite *mā...-š* negation in this corpus, illustrated in (59), and that this comes from one of the eighteenth- and nineteenth-century texts.⁴³

(59) ...bš'n m' yŕ'ybū-n-šy 'l n's
 so.that NEG blame.IMPF.3MSG-us-NEG the people
 ‘...so that people will not blame us.’

(C18th-19th Egyptian Judaeo-Arabic; Wagner 2007: 197)

⁴³ I transcribe the Hebrew consonantal script here directly without guessing at the intended quality of the unmarked short vowels.

Certainly this should not be taken to imply that bipartite negation only came to be a feature of Egyptian colloquial Arabic in this late period: it is already found in the fifteenth-century colloquial material in Ibn Sūdūn's *Nuzhat al-Nufūs wa-Muḍḥik al-ṣabūs*:

(60) **mā** maṯ-ak **ši** rūḥ

NEG with-you.MSG NEG soul

'You don't have a soul.'

(C15th Egyptian Arabic; Vrolijk 1998: 156)

This is the earliest example I have been able to find. The reason that bipartite negation is found in this and other Egyptian texts containing (usually satirical) colloquial passages (e.g. in Yūsuf al-Širbīnī's 17th-century *Hazz al-Quḥūf fī Šarḥ Abī Šādūf* – Davies 1981), but not in the letters of Wagner's Genizah corpus, must be due to the semi-formal nature of letter-writing. In fact, with almost no exceptions, the use of negation in Wagner's corpus is in conformity with the rules of Classical Arabic, or represents a hypercorrect interpretation of these rules (Wagner 2007: 168, 171). Note that this is stark contrast to many Spanish Arabic texts, particularly later texts from Granada (Corriente 1977: 143), where, as mentioned in section 2.3.1, one of the commonest negators is the entirely non-Classical *iš*. Hence, although we should be cautious in how we interpret the evidence of these texts precisely because of the likelihood of forms which are neither Classical nor vernacular, it is probably justified to take the lack of bipartite negation in all Spanish Arabic texts at face value without

drawing the same conclusion from the same lack in the rather different context of the letters from the Cairo Genizah.

In sum, then, the evidence reviewed so far does appear to support the hypothesis that Coptic played a role in the development of stage II negation in Arabic. There is, however, a final important obstacle to this hypothesis: the evidence of Jespersen's Cycle in certain Yemeni and Omani Arabic dialects. This is dealt with in the next section.

2.4.3 Jespersen's Cycle in the southern Arabian Peninsula

At first sight, the presence of stage II and III constructions in the south of the Arabian Peninsula represents something of a conundrum. In an earlier (2007) paper, following Obler (1990: 148), I made the somewhat unsatisfactory suggestion that Jespersen's Cycle could have spread here through sea trade contacts with Egypt. A somewhat more plausible scenario, some of whose difficulties are discussed at the end of this section, sees Jespersen's Cycle originating in the south of the Arabian Peninsula and being brought to north Africa by migrants of Yemeni origin. The third possibility, which I argue for here, sees the presence of Jespersen's Cycle in the south of the Arabian Peninsula as a development separate to Jespersen's Cycle in north Africa, albeit similarly contact-induced: this time involving the Modern South Arabian languages, which have themselves also undergone Jespersen's Cycle.

The Modern South Arabian languages are a group of closely related languages spoken primarily in the border region between Yemen and Oman. They form a branch of South Semitic, most closely related to the Ethiopian Semitic languages and

sufficiently different from Arabic as to be incomprehensible to monolingual speakers of the latter (Simeone-Senelle 1997: 378).

The most conservative Modern South Arabian language with respect to Jespersen's Cycle is Soqoṭri, spoken on the Yemeni island of Soqoṭra, where negation is with a stage I construction:

(61) ɔl fšek

NEG lunch.PRF.1SG

'I didn't eat lunch.'

(Soqoṭri; Simeone-Senelle 1997: 414)

The mainland varieties are either stage II or stage III. The former is represented here by Jibbāli, spoken in the mountains of Dhofar in south-western Oman:

(62) äxṭer ɔl kse mih her yofhəs ti-hum lɔʔ

caravan NEG find.PRF.3MPL water COMP boil.IMP.3MPL meat-their NEG

'The caravan didn't find water to boil their meat.'

(Jibbāli; Simeone-Senelle 1997: 413)

Note, as in Coptic, the late position of the postverbal element *lɔʔ*, which here follows even the embedded clause. *lɔʔ / laʔ* are the postverbal negators in all the stage II/III Modern South Arabian varieties, and in most of these the anaphoric negator 'no' is also *laʔ*. It seems likely, therefore, that *laʔ* 'no' was reanalysed as a clause-final negator in contexts where it was appended to the end of a sentence in either

declarative contexts ('I don't like that, no') or in tag questions ('You didn't like it, no?'), as suggested by Schwegler (1988) for the similar Brazilian Portuguese construction.

Among the stage II varieties, in Jibbāli at least, sentential negator *ʔ* appears to be fully grammaticalized as a negator rather than, for example, a negative polarity adverb, as it must occur in all negative contexts (with the exception of prohibitive sentences). Indeed, *ʔ* rather than *ɔ/* would appear to be the 'true' negator as the latter can appear pleonastically in non-negative downward-entailing contexts such as the VP-complement of verbs of fearing (on pleonastic or 'expletive' negation in the history of English see van der Wurff 1998; in Catalan, Espinal 1992; in West Flemish, Breitbarth and Haegeman 2009):

(63) *šeh ʔekəl ʔ yʔad*

he worry.PRF.3MSG EXPL.NEG go.IMP.F.3MSG

'He was worried about going.'

(Jibbāli; Johnstone 1981: 2)

Several Modern South Arabian varieties have in fact moved well into stage III of Jespersen's Cycle, maintaining the preverbal marker only in a restricted set of contexts. For example, Johnstone (1977: 2) remarks of Ḥarsūsi (also spoken in Dhofar) that, "*e/* [the preverbal negator], however, is common only in the double negative: *e/...w-e/* 'neither...nor'", which, as we have seen, is a common context crosslinguistically for conservatism with regard to Jespersen's Cycle. Otherwise we find a clause-final stage III construction:

(64) əkhōl əgəter la?

can.IMPF.1SG speak.IMPF.1SG NEG

‘I cannot speak.’

(Ḥarsūsi; Simeone-Senelle 1997: 414)

In view of these facts and the discussion above of the split distribution of stage II negative constructions in the Arabic dialects, there seems to be a reasonable *prima facie* case for saying that the stage II construction in Yemeni and Omani Arabic (as illustrated in (65) = (8)) was also triggered by contact, this time with earlier varieties of the Modern South Arabian languages.

(65) bih nās mā yiʕjib-hum-š aš-šāy

there.is people NEG please.IMPF.3MSG-them-NEG the-tea

‘There are people who don’t like tea.’ (Şanʕāni; Watson 1993: 261)

This being the case, the arguments for contact as the trigger for Jespersen’s Cycle in both the Arabic of north Africa and the south of the Arabian Peninsula now appear to be mutually reinforcing. If this is accepted then we also have doubled backing for the claim that contact can trigger syntactic changes in a language whose results do not perfectly replicate the structure transferred from the source language. Recall that in both Coptic and Modern South Arabian the postverbal negator typically occupies a clause-late position whereas in Arabic it is a verbal enclitic and thus precedes nominal and prepositional complements. At the same time the Arabic postverbal element cannot be described as a calque on either Coptic or Modern South Arabian, since the Arabic element is clearly derived from an indefinite pronoun,

whereas this appears not to be the case for its counterparts in either of the two proposed source languages. So we see that the negative constructions found in Coptic and Modern South Arabian are rather similar, while both differ noticeably from that found in the relevant varieties of Arabic. Nevertheless, this clause-late bipartite negative construction found in both source languages appears to have given rise to the Arabic verbal enclitic construction on two separate occasions and in two separate locations.

It may be objected at this point (Jonathan Owens, personal communication) that positing independent parallel development of the same construction in two different varieties of the same language ought to be avoided where a plausible case for historical connectedness can be made (the claimed non-independence of the Maltese stage II construction from that of mainland north African Arabic in section 2.3.1 was based on the same reasoning). Specifically in the present case, it is known that Arab tribes said to be of “Yemeni” origin were centrally involved in the conquest and settlement of Egypt (Owens 2003: 729).⁴⁴ Therefore, this line of reasoning would argue, we should see the presence of Jespersen’s Cycle in both north Africa and the southern Arabian Peninsula as linked by diffusion of a single innovation in the latter region.

Such a hypothesis has a good deal of initial plausibility, and it seems unlikely that any evidence could be produced that would force us to discount it altogether. Indeed,

⁴⁴ One should probably be cautious, however, in identifying the region inhabited by these tribes with the area occupied by present-day Republic of Yemen. As Owens (2003: 729 fn.26) points out: “Qaysites may be taken as roughly eponymous for Arabs of eastern Arabian peninsular origin, Yemenis for western [...] At this point in our study of the history of Arabic dialects, it would be premature to try to link these two broad groupings with specific dialect forms.”

if there was a stage II variant in the dialect of some of the Arab settlers in Egypt, then we would expect this variant to have been reinforced by the presence of the parallel Coptic stage II construction for the reasons given above. So accepting the possible accuracy of this hypothesis in no way requires us to abandon the hypothesis of a role for Coptic contact in the development of north African Arabic negation. However, positing a Yemeni – and denying a Coptic – origin for Jespersen’s Cycle in north Africa would appear to raise more questions than it offers answers, and, on its own, it would have to be viewed as a rather unsatisfactory explanation of the distribution of Jespersen’s Cycle among the Arabic dialects.

First of all, the Yemeni origin hypothesis involves stipulating that stage II negation had developed in at least some Yemeni variety prior to the Arab expansion out of the Arabian Peninsula. I am not aware of any independent evidence supporting such a stipulation. On the other hand, if this stipulation were accurate, and speakers of this variety were prominent enough in early Muslim society to be at the forefront of the aforementioned expansion, then one might expect some reference to stage II negation in the early Arab grammarians’ discussion of variation in Arabic dialects, particularly given their interest in variation concerning other aspects of negation (such as whether the predicate in copular sentences introduced with the negator *mā* is marked with nominative or accusative case; cf. Owens 1988: 26). There does not appear to be any such reference, though of course this lack of evidence in no way proves the non-existence of stage II negation at this stage.

Taking the Yemeni origin hypothesis further therefore, we are forced to ask once more, why do we end up with precisely the distribution of stage II negation among the Arabic dialects that we observe in Figure 1.2? Specifically, why should stage II

negation have taken root in Egypt but not in Iraq or Syria, where Yemeni tribes are also known to have been centrally involved in conquest and settlement? Concerning Syria in particular, Mad'aj (1988: 85) points out that:

“both early and modern historians confirm that the Yemenis formed the basis of the Muslim troops settling in Syria... In fact the nature of the settlement of the Arabian tribes in Syria was different from that in Iraq and Egypt. In Iraq and Egypt they confined themselves to the garrison towns [...], while in Syria they settled in towns and villages throughout various regions (*ajinād*).”

This being the case, the Yemeni origin hypothesis would have to introduce further stipulations to excuse the unexpected absence of Jespersen's Cycle in Syria. For example, Arabic came into intensive contact with the Aramaic varieties spoken by the local populations in Syria. Since Aramaic has only ever had stage I negation, and is structurally more similar to Arabic than Coptic is, it is conceivable that contact between the two favoured the conservative stage I construction and disfavoured the hypothetical innovative stage II construction. Again, this is a possibility that cannot be ruled out, but it seems at least as methodologically parsimonious to argue, as we have here, for independent parallel development under similar contact conditions, with no change in dialects that today lack a stage II construction, as it does to postulate a single development and explain its unexpected absence in various regions in an *ad hoc* fashion, or to leave it unexplained.⁴⁵

⁴⁵ Jonathan Owens (personal communication) also points out that stage II negation is absent in the Arabic of the Sudanic region, which is known to have been settled by migrants from

Finally, Jonathan Owens (personal communication) adduces another piece of evidence in favour of the Yemeni origin hypothesis, namely that the stage II construction is “morphophonologically specific and identical between Yemen and Egypt”. This refers to the effects that enclitic *-š* may have on the length of a preceding vowel and the position of stress in the word. However, it is not so clear that these effects are either specific to *-š* or identical in Yemeni and Egyptian Arabic. In Cairene Arabic, *-š* causes a preceding short vowel to lengthen and shifts stress to the right (66)a, but this is true of all consonant-initial enclitics (66)b.

(66) (a) *ma + 'gara + -š > maga'rāš* ‘it didn’t happen’ (NEG-happened-NEG)

(b) *is-'sana + di > issa'nādi* ‘this year’ (the-year-this)

(Woidich 2006: 34-35)

As far as stress is concerned, consonant-initial enclitics in Yemeni Arabic in general have the same effect, but there are exceptions which have to do with phonological properties of the root rather than the affix (see Watson 2002: 102-103

Upper Egypt at some point after 1225, indicating that the stage II construction cannot have spread throughout the whole of the Egyptian population by this time. However, these facts seem to me to be consistent with either the Yemeni or the Coptic origin hypothesis. Note in this connection that the people who brought Arabic to the Sudanic region were likely to have been nomadic pastoralists (Owens 2003: 722), and to this day the Arabic varieties spoken by (recently) nomadic Bedouin tend overwhelmingly not to feature a stage II construction, even within the region where the stage II construction is the norm amongst sedentary Arabic-speaking communities (cf. Caubet 2000-1: 79; de Jong 2000: 223, 477; Rosenhouse 1984: 515). *A fortiori*, we would not expect the stage II construction to have been a feature of the varieties spoken by nomadic Arabic-speaking tribes in Upper Egypt in the early second millennium either, even though it would, by hypothesis, already have been a feature of the variety spoken by the autochthonous sedentary population of the same region.

for details). Concerning vowel length, Watson (2002: 182) points out that, in the Yemeni variety she describes, “short morpheme-final vowels are rarely attested, and therefore the constraint on pre-suffixal short vowels applies vacuously.”

A further point of morphophonological non-identity concerns the resolution of a geminate consonant + *-š* cluster in the two varieties. In Yemeni Arabic there is degemination (67)a, whereas in Cairene there is epenthesis (67)b.

(67) (a) Yemeni: *mā + ḥabb + -š* > *mā ḥabš* ‘he didn’t like’ (NEG-liked-NEG)

(b) Cairene: *ma + ḥabb + -š* > *maḥabbiš* ‘he didn’t like’ (NEG-liked-NEG)

(Watson 2002: 210)

Thus the morphophonological properties of enclitic *-š* in Yemeni and Egyptian Arabic are similar, but not identical.

It seems, therefore, that while the Yemeni origin hypothesis cannot be ruled out at least as a contributing factor, the available evidence favours the Coptic contact hypothesis for the origin of Jespersen’s Cycle in north Africa. Note in this connection that other cases of independent parallel development of essentially identical constructions under similar language contact conditions can be pointed to. For example, the use of *bin* as a preverbal (non-progressive) past tense marker in a range of English-based creoles such as Tok Pisin (Verhaar 1995: 313; Austronesian substrate), Gullah (Turner 2002: 225; West African substrate) and Kriol/Australian Aboriginal English (Eades 1996: 134; Australian and Chinese substrate). Clearly here there is no temptation to say that the presence of this construction in all three of these creoles is the result of a single innovation within some earlier unrecorded

variety of English. Of course, the existence of the English past participle *been* was a necessary condition for the development of past tense *bin* in all three cases, but it cannot be seen as sufficient: standard English, for example, has not developed this construction. The sufficient condition for independent parallel development of this construction in each case can be assumed to be the properties of the tense-aspect systems of the respective substrate languages. It will be seen that this line of argument is directly parallel to the one I have advanced in favour of the independent parallel development of stage II negation in both north African and southern Peninsular Arabic: similar contact-induced pressure to develop a stage II construction exerted on similar Arabic varieties spoken in two separate regions leads to independent parallel developments.

In the following chapter I outline the model by means of which we will be able explain exactly how contact might have helped to bring about this change in both these cases.

2.5 Postscript: Jerusalem Domari

Before we close this chapter, a final possible case of the contact-induced spread of bipartite negation in this region should be mentioned. This comes from the highly endangered Jerusalem variety of the Indo-Aryan language Domari. Domari is spoken by the Dom people, a marginalized ethnic group of the Middle East and South Asia, and is closely related to, but distinct from, Romani, despite the comparable social status of the Roma people. So far the only variety of the language to have been described in any detail is the Jerusalem variety, sketched in Matras (1999, 2007). Here too we find a bipartite negative construction (68). It consists of a proclitic *n-*,

inherited from Proto-Indo-Aryan and ultimately Proto-Indo-European, and a stressed enclitic *-eʔ*, whose etymology is unknown, though Yaron Matras (personal communication) points out that the similarity with the numeral ‘one’ *ek* is suggestive of a derivation from some former indefinite item, conceivably an indefinite pronoun.

(68) *n-mangam-eʔ*

NEG-want.1SG-NEG

‘I don’t want.’

(Jerusalem Domari; Matras 2007: 152)

It is tempting to see this bipartite construction as a further case of contact-induced spread of Jespersen’s Cycle, this time from Palestinian Arabic. Unfortunately, however, the current state of our knowledge of other varieties of Domari makes it impossible to rule out an earlier innovation unconnected to that of Arabic. That said, there is sufficient record of at least the Turkish and Armenian varieties of the language to suggest that a bipartite construction is absent from these. The information we have on these varieties comes in the form of 46 sentences in the Bosha dialect of Armenia, published in Patkanoff (1907/8), and a glossary of the Zapari dialect of Turkey, published in Paspati (1870). In the first of these works there are three negative sentences, all of which contain a preverbal negator *na* and no other negative element. For example:

(69) charav na tharem

money NEG have.1SG

'I have don't have any money.' (Armenian Domari; Patkanoff 1907/8: 247)

In the second source we have a five-page entry on negation, with numerous examples listed of a number of forms: *na*, *nana*, *in*, *nanai*, *nasti*, *nastik*, *nanasti*, *nanastik*, *ne ne* and *ma* (Paspati 1870: 384-389). All of these are preverbal and no mention is made of any postverbal item (though, incidentally, it is interesting to note the extensive strengthening and reduplication of the preverbal negator itself here). This is clearly insufficient evidence to make strong claims of any sort, but it does seem to speak more in favour of bipartite negation being confined to Jerusalem Domari (and thus possibly the result of contact with Palestinian Arabic) than against it, thereby strengthening the impression of Jespersen's Cycle as an areal feature of north Africa and the southwestern Levant.

3. Modelling contact-induced change

3.1 Outline

This chapter outlines an approach to contact-induced change that will allow us to give an explicit account of the mechanisms underlying both the changes suggested in the previous chapter and, in principle, any hypothesized instance of contact-induced change. The chapter is structured as follows. In section 3.2 I give a critical overview of the field of contact linguistics, arguing that Van Coetsem's (1988, 2000) model compares favourably with other approaches in view of its conceptual clarity and focus on the psychology of individual speakers rather than social factors in speech communities. To be sure, a detailed understanding of the sociolinguistics of language contact is crucial for an account of what social conditions and social processes make contact-induced change possible in the first place; but an actual explanation of contact-induced change must involve an account of how contact leads to changes in the mental grammars of individual speakers. Van Coetsem's work offers a basis on which such an account can be built. Section 3.3 then discusses the fundamental ideas of Van Coetsem's framework in more detail. In section 3.4 I show how Van Coetsem's model can be adapted and improved in the light of empirical findings on the nature and consequences of bilingualism; we revisit the spread of stage II negation from Coptic to Arabic to Berber in the context of this discussion, as well as looking at two further case studies; and some key premises for understanding the mechanics of contact-induced change are established.

3.2 Problems with contact linguistics

There is, of course, already a well-established, if not yet mainstream, field of 'contact linguistics' which dates back at least to Schuchardt (1884), but which has been growing steadily since the pioneering work of Einar Haugen and Uriel Weinreich in the 1950s. However, as Van Coetsem (2000: 5) points out, "the study of language contact... remains largely stranded in an exploring or describing stage, and is still very much in search of its own theoretical premises." Van Coetsem's (1988, 2000) work has gone some way towards rectifying this state of affairs, as we will see shortly, but it has only recently begun to gain recognition (most prominently through Winford 2005, 2007).

Outside of the work of Van Coetsem and Winford, the theoretical debates in current contact linguistics remain centred on the putative limits of contact-induced change, imposed either by abstract properties of language conceived of as a structured entity relatively independent of the psychology of individual speakers (e.g. Zima 2007), or by sociolinguistic details of the contact situation (e.g. Mufwene 2007).

The key advantage of Van Coetsem's approach is that it provides us with a principled means of distinguishing between what have traditionally been known as 'borrowing' on the one hand and 'interference' (also referred to as 'substratum influence' or 'transfer') on the other. This it does by focussing on the bilingual speakers who are the agents of contact-induced change and on the different types of agentivity they employ, depending on whether they are dominant (psycholinguistically rather than socially speaking) in the source language (SL) or the recipient language (RL), that is, the influencing and the influenced language respectively (see section 3.3 for further details).

Many of the long-standing debates over the limits of contact-induced change have remained unresolved in part due to the lack of such a principled means of classifying changes into instances of borrowing or interference. Hence, for example, Thomason and Kaufman (1988: 49-50) classify Ma'á⁴⁶ as having undergone “massive grammatical replacement”, as a result of large-scale borrowing of SL grammar on the part of the RL. This classification falls out naturally from their sociolinguistic understanding of what constitutes borrowing: “Borrowing is the incorporation of foreign features into a group’s native language by speakers of that language: the native language is maintained but changed by the addition of the incorporated features” (Thomason and Kaufman 1988: 37). Here there is no regard for the agentivity involved in contact-induced change: the fact that an RL has not been abandoned by its original speakers automatically means that any changes that it undergoes as a result of contact with another language count as instances of borrowing, irrespective of whether the changes were introduced by speakers dominant in the RL or the SL. Put another way: they see an automatic link between language maintenance and borrowing on the one hand, and language shift and interference/substratum influence on the other.

This allows Thomason and Kaufman to point to a number of examples which count as instances of large-scale grammatical borrowing in their terms (that is, where a maintained language shares its inflectional morphology, word order and other grammatical features with a contact language) in support of their claim that “[a]s far as the strictly linguistic possibilities go, any linguistic feature can be transferred from

⁴⁶ Ma'á is a so-called ‘mixed’ or ‘intertwined’ language spoken in Tanzania, whose morphosyntax is largely identical to that of the Bantu language Mbugu, while its lexicon is essentially Cushitic (Thomason 1983).

any language to any other language” (1988: 14). Now, according to their definition of borrowing given above this claim is unexceptionable. The fact that many have taken exception to it (see e.g. Silva-Corvalán 1998, King 2000, Myers-Scotton 2002), without addressing the lack of consensus on the definitions of the terms involved, points up another of the major problems in current contact linguistics: ambiguous, inconsistent and vague use of terminology.

Winford (2005: 374) is one of the few who have drawn attention to this confusing state of affairs:

“‘interference via shift’ has also been referred to as ‘substratum influence’ especially in the role of Creole formation, and as ‘transfer’, in the context of SLA [= second language acquisition]. Some scholars use the term ‘interference’ to refer to any type of crosslinguistic influence, including borrowing, while others use ‘transfer’ in the same broad sense.”

To this we could add Weinreich’s distinction between what he calls ‘interference’ and ‘borrowing’, which is wholly orthogonal to the distinction that most people make between these terms. For Weinreich (1953: 1), ‘interference’ “implies the rearrangement of patterns that result[s] from the introduction of foreign elements into the more highly structured domains of language,” whereas his ‘borrowing’ refers to the introduction of foreign elements which do not affect the ‘system’ or the ‘structure’ of the RL.

Others, perhaps understandably, have preferred to develop entirely their own terminology to describe the processes and outcomes of language contact: compare,

for example, Johanson's (2002) term 'code copying'. In fact, we find that there are almost as many new technical terms or alternative uses of existing terminology as there are linguists working on the theory of language contact. This is probably both the result of, and a continuing contributing factor to, the atomization of the field, whereby each researcher works in his or her own individual framework, lacking a coherent collaborative research programme as long as the study of language contact remains "in search of its own theoretical premises", as Van Coetsem (2000: 5) observes. As stated above, the framework developed by Van Coetsem (1988, 2000) and taken up by Winford (2005, 2007) goes some distance towards establishing some of these basic theoretical premises, which we examine in more detail in the following section.

3.3 Van Coetsem's model

Over the course of two monographs (1988, 2000), Frans van Coetsem (lived 1919-2002) presents his model of contact-induced change in extensive detail. This model, having been largely overlooked for a time,⁴⁷ has recently been defended and applied to some of the more contentious cases of contact-induced change in the literature by Winford (2005, 2007). It is not my intention here to review these works in full. Instead, in this section I present the essential elements of the model, and in section 3.4 discuss some of the ways it might be adapted and improved upon in the light relevant empirical research on bilingualism. This provides the basis for the discussion of the mechanics of contact-induced change given in chapter 4. It is my position that, although there are aspects of Van Coetsem's framework which are dispensable, the

⁴⁷ Exceptions that Winford (2005: 376) points out are Guy (1990) and Smits (1998).

key distinction between RL and SL agentivity is crucial to any account of contact-induced grammatical change which aims at psychological reality.

The introduction of material from one language into another is called ‘transfer’ by Van Coetsem. Transfer is, by definition, always from the ‘source language’ (SL) to the ‘recipient language’ (RL). Two main transfer types may then be distinguished: ‘borrowing’ and ‘imposition’. Borrowing designates much the same thing here as it does in general usage, while imposition aligns broadly with the notion of ‘interference’ or ‘substratum influence’, however, here these terms are given precise definitions. It is important to note that in this framework borrowing and imposition are not mechanisms of change in their own right, but merely the major types of change that occur in contact situations as a result of psycholinguistic processes at work in the minds of bilingual speakers of the languages involved (Winford 2007: 26).⁴⁸ These speakers are the ‘agents’ of contact-induced change.

Ordinarily the agents of change will be ‘dominant’ in one of the languages involved in a contact-induced change. Crucially, ‘dominance’ here is a purely psycholinguistic, not social, notion. In section 3.4 I argue that a bilingual speaker’s dominant language is equivalent to her first, or ‘native’, language (L1), that is, the language she acquired from birth (leaving aside for the moment the question of ‘balanced’ bilinguals). However, Van Coetsem (2000: 52) argues that “a speaker may become more proficient in a subsequently acquired language than in his native language (which comes first in time)... Nativeness can thus be overruled by linguistic

⁴⁸ I use ‘bilingual’ here and throughout as shorthand for ‘bilingual and/or multilingual’.

dominance,” (see also Smits 1998: 378).⁴⁹ When an agent of change is dominant in the RL, a change she makes to the RL is said to occur under ‘RL agentivity’, which is the equivalent of borrowing. When on the other hand an agent of change is dominant in the SL, a change she makes to the RL is said to occur under ‘SL agentivity’, which is equivalent to imposition. Note that on this definition of agentivity there is the potential for one or many bilingual speakers to employ each type of agentivity on two (or more) different languages in a single contact situation: RL agentivity on the language in which they are dominant, and SL agentivity on the language in which they are non-dominant.

With this overview of the two transfer types in place we are now in a position to illustrate with some uncontroversial examples. One of the most salient effects of borrowing under RL agentivity is the transfer of lexical items (sound-meaning pairings) from the SL to the RL. Arabic, for example, has a number of loan words of European origin whose basic form betrays their foreign origin, but which have been fully integrated into the Arabic phonological and morphological systems. For instance, Arabic nouns such as *film* ‘film’, *sigāra* ‘cigarette’, and *faylasūf* ‘philosopher’ are, to a naïve native speaker, indistinguishable from their non-loaned equivalents in that their basic meaning is associated with the consonants, or ‘root’, while a grammatical category such as number is expressed by the addition, subtraction, or change in quality of vowels: *aflām* ‘films’, *sagāʔir* ‘cigarettes’, *falāsifa* ‘philosophers’. In the same way, words such as Graeco-Latin *telephone* and

⁴⁹ Whether proficiency is a useful notion from a psycholinguistic point of view will be discussed in section 3.4, but suffice it to say for the moment that a distinction between a speaker’s L1 and the language in which she can be said to be most proficient does not arise in the majority of cases.

television have been integrated into Arabic's derivational morphological system to produce *talfana* 'to telephone' and *talfaza* 'to televise'. It is clear that these words were introduced into Arabic by native speakers, that is, through borrowing under RL agentivity, in order to provide labels which were lacking for recently introduced concepts, and then adapted to fit Arabic morphophonology.

The most obvious and apparently widespread manifestation of imposition under SL agentivity is in the realm of phonology. The extent to which bilingual speakers who are dominant in the SL tend to impose the phonology of the SL on the RL is such that this phenomenon is well-known in popular culture as 'foreign accent'. Concrete examples are, for instance, cases where a speaker who is dominant in French fails to aspirate initial stops in English, so pronouncing a word like *pit* as [pit] rather than [p^hit] (Van Coetsem 1988: 11), or where a speaker who is dominant in Greek substitutes a voiceless alveolar fricative for the English voiceless postalveolar fricative, which is lacking in the phonemic inventory of standard Greek, and so pronounces a word like *shops* [sɔps] instead of [ʃɔps].

Imposition clearly has syntactic manifestations too, for instance in changes to the argument structure of verbs as in (70), where a speaker who is dominant in English assumes that the thematic role of SOURCE for the French verb *emprunter* 'to borrow' is marked with the preposition *de* rather than the standard French *à* on the model of English *to borrow from somebody*, or (71), where a German-dominant speaker assumes that bound reciprocal anaphors in English are identical in form to accusative first and second person plural pronominals, as is the case in German:

(70) *Je l'ai emprunté de mon ami.

I it-AUX.PRF.1SG borrow.PTCP from my friend

'I borrowed it from my friend.'

(71) (a) *We will see us tomorrow.

(b) Wir sehen uns morgen.

we see us/each other tomorrow

'We will see each other tomorrow.'

Of course, historical linguistic research tends primarily to focus on changes that are observed at the level of the speech community, rather than on sporadic examples of change in individual idiolects, and the present thesis is no exception. However, as Thomason (2007: 25) points out, echoing Weinreich et al. (1968): "Every change in a community's language of course involves two steps – first an innovation in the speech of one or more speakers, then spread of the innovation to the rest of the community's speakers" (see Hale 1998 for more discussion on this point).

The fact is, while lexical transfer at least is clearly identifiable as such even after its results have become integrated into the RL at the level of the speech community, innovations in the RL involving syntactic transfer are most clearly seen before they spread throughout the speech community, as in (70) and (71). Once they have spread, their effects on the RL can be hard to distinguish from internal innovations, all else being equal (and this is perhaps another cause of the traditional *a priori* preference for purely internal accounts of change over ones involving contact; cf. section 2.1). However, given sufficient data concerning the grammars and the sociolinguistic situation of the languages in contact prior to the relevant innovations,

we will usually be able to say with a fair degree of confidence whether or not those innovations are contact-induced, and, if so, whether they are due to RL or SL agentivity. Furthermore, it could conceivably be, as Van Coetsem (1988, 2000) and others have argued, that borrowing and imposition characteristically affect different domains of language, and that therefore a simple inspection of the types of elements or structures that have been transferred will allow us to say which of the transfer types was responsible in each case. However, we will see in the following section that there is reason to doubt the success of the inspection method in accounting for all types of contact-induced syntactic change.

3.4 Expanding and revising Van Coetsem's model

At this point several questions arise concerning the basic elements of Van Coetsem's model introduced so far, which need to be dealt with prior to discussion of the actual mechanisms underlying contact-induced grammatical change. First, how exactly are we to understand this notion of 'dominance' on which the distinction between RL and SL agentivity rests? Second, are there any distinctive characteristics of borrowing and imposition, as characterized above, such that we can distinguish them via their results? And finally, do borrowing and imposition exhaust the possible types of contact-induced change?

3.4.1 Dominance

As indicated above, Van Coetsem (2000) equates linguistic dominance with 'proficiency' (which he leaves undefined) and makes the point that a speaker may, under certain circumstances, become more proficient in a second language (L2) than

in her L1 (2000: 52). If proficiency is calculated by means of fluency tests, vocabulary tests, tests of written style and the like, then it seems likely that many people do in fact become more proficient in an L2 than in their L1. However, it is not at all clear that this has any meaningful implications for the study of the role of bilingualism in contact-induced change in grammars. In particular, that this may be the case does not entail that the L1s of these individuals have been impacted in any way in terms of their underlying grammatical competence, as distinct from more ephemeral performance or processing considerations.

It is uncontroversial that linguistic performance in an L1 can be noticeably affected by prolonged exposure to and use of an L2 (and concomitant lack of exposure to and disuse of the L1), but whether the same can be said for adult L1 competence is an empirical question, and one which, in contrast to questions of proficiency, is central to an understanding of the cognitive mechanisms underlying contact-induced change, as it has a crucial bearing on how borrowing takes place.

In what follows, I use the term 'attrition' as short-hand for 'some bilingualism-induced alteration in a speaker's L1 competence and/or performance'. Importantly, 'attrition' here refers not only to deterioration, but to any alteration at all, including those best thought of as the addition of new options or rules.⁵⁰ Unfortunately there has not been a great deal of empirical work conducted specifically in order to explore the question of whether competence as well as performance can be attrited in this way, not least, perhaps, because separating competence from performance in empirical data can be highly problematic (cf. Altenberg and Vago 2004).

⁵⁰ Compare Seliger (1996: 614 fn.6): "Language attrition is, in fact, a creative and often additive process."

What, then, is the most reasonable hypothesis to make concerning this question, until such time as evidence can be produced that will falsify it? Sharwood Smith and Van Buren (1991: 20) take the position that, “it does seem reasonable to suggest, at least as a working hypothesis, that the attrition of competence may be triggered by changes in the learner’s perception of the basic structure of his or her L1 grammar and not just by a tendency to ease the processing burden of an underused L1.” However, it would appear that few others who have engaged with this issue in the recent L1 attrition literature favour this point of view. Instead, researchers in this field have tended to assume (following Chomsky 1980; e.g., Köpke 2004, Flores 2007 and Tsimpli et al. 2004) that, upon maturation, a human being’s Language Acquisition Device reaches a ‘steady state’, and the grammar that instantiates this steady state does not undergo any significant alteration in non-pathological individuals during adulthood, whether or not an L2 is subsequently acquired.

For example, in her monograph on L1 attrition among German Jews in Anglophone countries, Schmid (2002: 18) finds concerning both morphology and syntax that, “the question of whether [...] the actual knowledge of the L1 can become deteriorated [...] has not been conclusively resolved, but evidence overwhelmingly points towards what difficulties there are being only temporary.” This position is strengthened by a series of reports (Van As 1962, Fromm 1970, Footnick 2007) in which subjects who professed to having completely lost the language of their parents displayed native-like ability in both comprehension and production of that language whilst under hypnosis. It is natural to suppose that what holds for these extreme cases also holds for commoner, less extreme cases of L1 attrition: the appearance of fully-fledged loss of linguistic knowledge, considered as such by the subjects

concerned, may in fact be merely the result of temporary difficulties with access and retrieval.

The findings of a study by Altenberg (1991), designed to investigate the effects of L1 attrition on syntax, illustrate this point. The study reports on a series of tests conducted on two native speakers of German who had been living in the United States for over forty years and who spoke fluent but non-native English. In an attempt to minimize processing and performance factors, all of the tests consisted of untimed acceptability judgment tasks. The most relevant of these for the purposes of the current discussion was a syntactic judgment task in which the subjects were presented with various German and English sentences and asked to judge their acceptability on a five point scale. The sentences were of four types: those whose word order is the same in English and German (72), those whose word order is grammatical in English but ungrammatical in German (73), grammatical in German but not in English (74), and finally, grammatical in neither German nor English (75):

(72) Das ist leicht zu machen.

that is easy to do

(73) *Barbara kann lesen hindi.

Barbara can read Hindi

(74) Barbara kann hindi lesen.

Barbara can Hindi read

(75) *Sie wohnt in einem Haus blauen.

She lives in a house blue

(Altenberg 1991: 192-193)

The overall results showed that both subjects had a firm grasp of word order in both German and English. However, for several German sentences whose word order is ungrammatical in standard German (some of them were acceptable in some varieties of colloquial German), one or both subjects judged them to be fairly acceptable, for example:

(76) *?Der Mann, dessen Gepäck steht da, kommt gleich zurück.

The man whose luggage is standing there is coming right back.

(Altenberg 1991: 194)

Interestingly, in follow-up interviews conducted several weeks after the main experiment, “both subjects stated that all of these [...] structures [...] were fairly unacceptable in German and expressed surprise at their own responses on the judgment task regarding these structures” (Altenberg 1991: 196). Why, then, were these marginal or ungrammatical sentences judged acceptable first time around? In the light of the above discussion, a reasonable answer seems to be that this was not because the sentences were grammatical in the subjects’ colloquial variety in its original unattrited state, nor because their underlying grammatical competence had been attrited and brought into line with English, but because processing difficulties caused by strong activation of the relevant English structures in the experimental setting resulted in performance-based attrition (despite the experimenter’s efforts to test competence rather than performance; cf. Schütze (1996), Cowart (1997) on the inherent involvement of performance factors in acceptability judgments). When this

source of processing difficulties was removed in the follow up interviews, the subjects' judgments were able to reflect more closely their underlying, unattired, competence.

Thus, until such time as empirical evidence to the contrary is produced, it seems reasonable to proceed on the uncontroversial assumption that in non-pathological individuals L1 attrition is restricted to the level of performance and does not alter competence. This has two important consequences. Firstly, if we take the innatist view of syntactic change as change in mental grammars, then what we have in L1 attrition is not yet syntactic change. Rather, as we will see in chapter 4, L1 attrition is best viewed as a source of perturbation to the primary linguistic data (PLD) on the basis of which children acquire their L1, and thus an important input to change, rather than change itself. The second consequence would appear to be that dominance in an L1 cannot be replaced by dominance in an L2. This point is reinforced by a consideration of the nature of knowledge of an L2 versus knowledge of an L1.

As is well known, the process of acquiring knowledge of an L2, particularly as an adult, is typically much longer, more laborious, and ultimately less successful than that of acquiring an L1.⁵¹ While there is a considerable body of literature that seeks to model L2 performance (henceforth referred to by its usual name in the L2 acquisition

⁵¹ It is also well known that this imbalance between L1 and L2 is greatly reduced when L2 acquisition begins in childhood. Nevertheless, it appears to be the case that even child L2 learners exhibit some differences to L1 acquirers, particularly in the domain of inflection (cf. Wexler 1998; Blom, Polišenská and Weerman 2006; Meisel 2008, 2009), unless those children are very young (perhaps less than four or five years old at most). This raises the question, which cannot be dealt with here, of how to distinguish between the simultaneous acquisition of two first languages and L1 and L2 acquisition in young children. In the absence of a present means of teasing out these distinctions, the rest of this thesis will focus on adult L2 acquisition, with no implication intended that child L2 acquisition is excluded as a possible source of contact-induced change.

literature: 'interlanguage') using the tools of generative grammar as developed to describe L1 competence, it is far from clear that either the process of acquisition of an L2, or the character of the knowledge of that L2 that is ultimately attained, are the same as we find in L1 acquisition. Instead, the observed discrepancies between L1 and L2 acquisition receive a principled explanation if we assume something like Bley-Vroman's (1989) Fundamental Difference Hypothesis.

This idea, which has been echoed in much contemporaneous and subsequent work (see e.g. Clahsen and Muysken 1989, Eckmann 1996, Schachter 1996, Meisel 1997), essentially states that L1s are acquired on the basis of an innate, domain-specific Universal Grammar (UG), whereas L2s are acquired by means of a range of rather more general learning strategies.⁵² That is, while 'poverty of the stimulus' arguments imply the necessity of UG in order to constrain the set of hypotheses

⁵² The Fundamental Difference Hypothesis as originally formulated was always intended as a robust and provocative statement of a particular approach to understanding L2 acquisition, rather than a definitive statement of the way things must be. In the light of developments in linguistic theory in the last twenty years, Bley-Vroman (2009) has recently revisited the hypothesis, arguing that a significant reformulation is required. The *explanandum* – that knowledge of an L2, unlike knowledge of an L1, is typically unreliable and non-convergent – is as blatantly the case as it ever was. The *explanans*, however, can no longer rely on a richly structured UG and purely domain-specific acquisition processes, since these have been rejected in contemporary thinking, from both the mainstream generative (Minimalist) and competing (Construction Grammar, Simpler Syntax) perspectives. Instead, Bley-Vroman (2009) now suggests that 'patches' (Morgan 1972, McCawley 1988) or 'viruses' (Lasnik and Sobin 2000), which undoubtedly feature in the periphery of L1 knowledge (e.g., in the understanding and use of the prescriptive rules of a prestige variety), and which have the properties of unreliability (speakers are unsure whether patches are grammatical) and nonconvergence (agreement on the grammaticality of a given patch is the exception rather than the rule), are at the core of L2 acquisition, explaining why L2 knowledge has these properties. This proposal seems to be on the right track, but we will not pursue it further here. The crucial point for present purposes is that the basic observation remains, whatever its explanation: speakers' knowledge of an L2 is of a different character to their knowledge of their L1.

children must make concerning the grammar of their L1s, what has been called 'the logical problem of L2 acquisition' (how it is that interlanguage has characteristics which transcend the input) requires only the availability of adults' general problem-solving and analytical abilities along with knowledge of their L1s.

Support for this idea has come from a number of empirical studies. For example, in a study of 57 L1 Hungarian immigrants to the US, DeKeyser (2000) found that a native-like score on an English syntactic acceptability judgment task was a statistically significant predictor of high scores on a general verbal analytical ability task (in Hungarian). This is in contrast to L1 speakers, of course, where knowledge of the grammar of the L1 appears to be independent of other cognitive abilities. Here the clear implication is that verbal analytical abilities are crucial to the eventual target-like acquisition of an L2, while they are largely irrelevant to the acquisition of an L1.

As regards the underlying character of knowledge of an L2, Johnson et al. (1996) conducted a study comparing a group of L1 English speakers with a group of L1 Chinese adult immigrants to the US, who had been resident in the US for a sufficiently long time for their L2 English development to have plateaued. Each group completed an acceptability judgment task on two occasions with an intervening period of three weeks. While the native speakers' answers exhibited a predictably high degree of consistency, the answers of the L2 group showed a significant degree of variability between the two tests. The conclusion that Johnson et al. draw from this is that knowledge of an L2 may not be the same type of formal object as knowledge of an L1, in that the former appears to be inherently indeterminate, in marked contrast to the latter.

If we accept that, in the majority of cases at least, there is a substantive difference between the nature of an L1 and an L2, and that L1 competence does not, generally speaking, undergo attrition, then where does this leave the notion of dominance and the idea that a speaker's L2 can become dominant and her L1 non-dominant?

Two conflicting ways that the notion of dominance can be defined such that it is psycholinguistically more precise suggest themselves as follows.

Definition 1) A speaker's dominant language is whichever of her L1, L2, L3, etc., is most accessible at any given time.

Definition 2) A speaker's dominant language is her L1. Any other language subsequently acquired is necessarily non-dominant.

It is clear that for many bilinguals the accessibility of their L1s and L2s is shifting constantly. Under the first definition of dominance, then, an L2 can certainly become dominant with respect to an L1. However, this definition would appear to have the unwelcome consequence that transfer under RL agentivity into one's L1 – that is, the most intuitively obvious type of borrowing – becomes logically impossible. If, at a given time, a speaker transfers some feature of her L2 into her L1 performance, then presumably this is at least partly because that feature is more accessible to her at that time than its L1 counterpart. But under the first definition of dominance, this would then be a case of imposition under SL agentivity – transfer from the dominant to the non-dominant language. For this reason, and given what appear to be reasonable assumptions concerning the respective natures of L1s and L2s, it seems

more satisfactory to adopt the second definition of dominance. Under this definition an L1 is necessarily always dominant with respect to an L2, regardless of its accessibility at any given time. As such, it will occasionally be useful in what follows to make use of the terms L1 and L2 agentivity alongside the now familiar RL and SL agentivity, the two pairs being largely equivalent though with different emphases.⁵³ A welcome consequence of adopting the second definition of dominance is that we now have a principled basis for explaining why borrowing and imposition each seem to have characteristically different consequences for an RL: to the extent that this is the case, it can be attributed, at least in part, to the very different processes of L2 acquisition and L1 performance attrition. This issue is discussed in detail in section 3.4.2.

Before we leave the topic of dominance, however, a word is required on 'balanced' bilinguals, that is, those speakers who have undergone the simultaneous acquisition of two first languages (2L1 acquisition). Clearly in the case of such balanced bilinguals neither of the two L1s is dominant and so the distinction between borrowing and imposition breaks down. Van Coetsem (2000) acknowledges this and talks of "neutralization". He goes on to suggest (2000: 86) that neutralization permits the "free transfer" between the two L1s of elements from any linguistic domain. Whether this is what in fact occurs is again an empirical question, however, and one which has received some attention in the literature.

⁵³ RL/SL agentivity places greater emphasis on the languages to and from which material is transferred, while L1/L2 agentivity places greater emphasis on the psycholinguistic processes involved in L2 acquisition and L1 attrition. This distinction becomes especially important when we consider changes that are contact-induced and occur under L2 agentivity, but do not involve the transfer of anything from a source language.

Several researchers of (bilingual) child language acquisition, such as Döpke (1998), Müller (1998) and Müller and Hulk (2001), have claimed that children simultaneously acquiring two L1s transfer syntax between the two grammars they are acquiring.⁵⁴ Others disagree. Meisel (2001: 39-40), for example, argues that “the *Interdependent Development Hypothesis* [is] less convincing than the *Autonomous Development Hypothesis*”, that is, it is less likely that there is syntactic transfer in the course of 2L1 acquisition (at the level of competence rather than merely performance) than that there is not.

In support of this view, Meisel (2001) gives various arguments against the claims of the above-mentioned authors, including, importantly, what he calls the parameter resetting problem: given that balanced bilinguals by definition eventually achieve the same grammatical competence in each of their L1s as their monolingual counterparts, a period during acquisition where they set certain parameters to the ‘wrong’ values due to influence from the other language must necessarily be followed by a period in which these parameters are reset to the ‘correct’ value (Meisel 2001: 37). All of the above-mentioned authors couch their arguments in generative terms, but the notion of parameter resetting goes against standard generative assumptions about acquisition (Tsimplici and Roussou 1991, Meisel 1995, Fodor 1998). The opposite view is argued for by Müller (1998), for example, who claims that German

⁵⁴ Note that this does not entail a claim of support for Volterra and Taeschner’s (1978) Fusion Hypothesis. This idea, which has largely been abandoned in recent work on bilingualism, states that bilingual children’s mental representation of the two languages they are acquiring consists of a single unified system. In contrast to this, there is now widespread consensus that children raised in a bilingual environment differentiate the grammatical systems of the languages they are exposed to from the outset (see Meisel 2001, 2004 for a summary of the relevant research).

target-deviant utterances, such as (77), produced by a French/German bilingual child, in which the verb appears in medial rather than final position, are indicative of transfer.

(77) Sagen wir mal, dass das is ein Baum.

say we PRT that this is a tree

'Let's say that this is a tree.' (Müller's 6b.)

However, exactly the same kinds of errors of verb placement occur in the speech of monolingual German children (Gawlitzeck-Maiwald, Tracy and Fritzenschaft 1992), as Müller (1998: 160) acknowledges. She claims, though, that there is, "a quantitative difference between the two acquisition types: whereas the error types presented here are frequently encountered in bilingual language development, they seem to represent rather an exception in monolingual language acquisition." (Müller 1998: 160). Unfortunately Müller (1998) gives no statistics to substantiate this assertion (cf. de Houwer 1998), making its validity rather hard to assess. Let us, however, assume that this quantitative difference is genuine – this still does not mean that this is certain evidence of transfer at the level of competence/parameter settings. Rather, just as we have seen in the discussion of L1 attrition, it is more likely that errors such as the one in (77) are the result of processing difficulties which are exacerbated in cases of bilingualism. That is, if they represent transfer at all, it is at the level of performance, rather than competence.

The idea of a certain degree of availability, during processing, of lexicon and syntax from both a bilingual's languages, and the recruiting of elements of one

language into the frame of the other as a performance-based relief strategy, is uncontroversial (cf. Genesee 1989, Meisel 1989, and, indeed, the entire literature on code-switching). Thus it seems preferable, in the absence of really convincing evidence to the contrary, to assume with Meisel (2001) that in 2L1 acquisition, as in L1 attrition, transfer may take place at the level of performance, but not at the level of competence. What we see, then, is that apparently neither L1 attrition nor 2L1 acquisition result, in and of themselves, in a change in mental grammars. Rather, as will be argued in chapter 4, it is children acquiring L1s (whether monolingually or otherwise) who bring about grammar change in situations of contact, simply because of the alteration to their PLD brought about by non-balanced bilingual speakers in their environment.

3.4.2 The 'stability gradient'

In this section we deal with the question of to what extent the linguistic results of borrowing and imposition differ in characteristic or predictable ways. This is important to the present discussion for two reasons. Firstly, both Van Coetsem (1988, 2000) and, following him, Winford (2005) argue that borrowing and imposition do indeed have characteristically different results and attribute this to the so-called 'stability gradient'. The importance of this idea to Van Coetsem's framework is such that it merits careful scrutiny here. Secondly, if borrowing and imposition can be shown to have characteristically different results, then this suggests that one should be able to reconstruct which type of transfer was operative in a particular change simply by inspecting the linguistic results of that change. The question of whether, or to what extent, this is in fact the case is clearly of central importance to this kind of

framework. I argue that the general principle – that borrowing and imposition are to some extent associated with different types of change – is sound, but that these differences are best understood in the context of the approach to dominance developed above, whereby borrowing and imposition are rigidly associated with L1 and L2 agentivity respectively. Moreover, I will try to show that the differences between the results of borrowing and imposition are not quite as Van Coetsem and Winford would have them, particularly as regards the transfer of syntax.

The stability gradient is based on the idea that “language does not offer the same degree of stability in all its parts, in particular [...] there are differences in stability among language domains, namely among vocabulary, phonology and grammar (morphology and syntax).” (Van Coetsem 1988: 25). It is claimed that these differences essentially amount to the fact that “phonology and grammar in general show greater stability than vocabulary” (1988: 26), and that in both borrowing and imposition it is the more stable domains that are preserved.

Winford (2005: 377), who claims a strong analytical power for the idea of the stability gradient, tells us that “this explains why borrowing tends to be mostly lexical, and does not usually affect the RL grammar, while imposition, on the other hand, tends to do so.” This would seem to be a fairly uncontroversial, common sense statement, and this idea certainly does not originate with Van Coetsem. Haugen (1950: 224), for example, invokes the very similar ‘scale of adoptability’, although this refers only to borrowing rather than to transfer in general (cf. Van Coetsem 1988: 165, fn. 33), Moravcsik (1978) proposes several universal constraints on borrowing which entail much the same thing, and Thomason and Kaufman (1988: 50) have a “scale of interference”, although they stress that this is predicated upon

sociolinguistic factors such as the intensity of contact, and explicitly deny, as we have seen above (section 3.2), that there are absolute linguistic constraints on transfer in any domain.

However, there are problems with this idea on several levels. First of all, as Smits (1998: 388, fn. 5) correctly points out:

“The fact is that we do not have a method at our disposal to ‘weigh’ the degrees to which the respective linguistic components have been affected by interference. In other words, we actually cannot measure exactly whether lexical interference is more prominent or less prominent than phonological/phonetic/syntactic/morphologic interference.”

It remains to be seen whether such a method of empirically verifying claims of the relative prominence of transfer in different linguistic domains can be found, or whether, in fact, the qualitative differences between the different linguistic domains makes this a logical impossibility.

Another important issue for the idea of the stability gradient concerns the accuracy of the generalization which it is invoked to explain. This is crucial for the arguments of the rest of this chapter, so it is worth tackling in some detail here. At least one of the claims associated with the stability gradient does appear to be testable: namely, that a given linguistic domain will be involved more in one of the two major transfer types than it will in the other. So, for example, it is said that phonology is imposed more than it is borrowed, while lexicon is borrowed more than it is imposed. These claims seem likely to be true, and I will not discuss them further

here. Let us instead focus on Winford's (2005: 377) claim, cited above, that "borrowing [...] does not usually affect the RL grammar, while imposition, on the other hand, tends to do so". In the next two subsections I attempt to show that this statement, while it is not outright false as it stands, is rather misleading. For example, in his discussion of the heavily Turkish-influenced dialects of Cappadocian Greek Winford (2005: 407) asserts that "the transfer of Turkish rules of [...] morphosyntax [and] word order, especially in copular and interrogative constructions, are symptomatic more of imposition than of borrowing." We will see that the general assumption that the transfer of abstract elements of syntactic structure is probably imposition and not borrowing is unjustified and in need of modification.

3.4.2.1 Grammatical borrowing

'Grammatical borrowing', for the purposes of this debate, is used synonymously with 'structural borrowing' to refer both to the borrowing of bound and unbound morphemes (form as well as function – what Matras and Sakel 2007 call 'matter replication'), and to the borrowing of abstract syntactic patterns or structure (what Matras and Sakel 2007 call 'pattern replication'). Winford (2005) only overtly discusses the possibilities of the former variety of structural borrowing, which he concedes is possible, but "only when the languages involved are typologically very similar, allowing for the substitution of an RL morpheme by a close counterpart in the SL" (Winford 2005: 387). Harris and Campbell (1995: 124), on the other hand, claim that there are plenty of counterexamples to this supposed restriction. However, given that 'typological similarity' or 'structural compatibility' is another notion for which there

is no established metric, or even an agreed-upon definition, there appears to be no rational basis on which to settle this argument.

Let us instead focus here on the latter type of grammatical borrowing (pattern replication). My contention will be that the Arabic-influenced development of stage II negation in Berber represents an example of this type of grammatical borrowing. I then give a second likely example from a dialect of Konkani (Indo-Aryan) described by Nadkarni (1975). How common change due to the borrowing (and indeed imposition) of syntactic structure is beyond these examples remains an empirical question.

I have already made the case in section 2.3 for contact with the relevant Arabic varieties being implicated in the development of stage II negation in the Berber varieties that have it. In order to give an explicit account of how it was that contact led to this innovation in this (and any other) case, it is essential to establish as far as possible which of the two types of agentivity was involved. What we can infer about the sociolinguistic situation in north Africa from the time of the Arab conquest until the present day points to this change having been primarily the result of borrowing (under RL agentivity), rather than imposition (under SL agentivity). It should be stressed, however, that some role for imposition cannot be entirely ruled out (just as the Berber and Arabic bipartite constructions having developed independently cannot be entirely ruled out, even given the evidence suggesting contact-induced change). As such, when we go on in chapter 4 to examine the mechanisms by which this change took place, we are necessarily modelling an idealization of what occurred rather than certain facts in all their real-life complexity. Nevertheless, it seems useful

to offer an explicit (and falsifiable) account of this and other historical changes, even if that account is based on reasonable assumptions rather than certainties.

The reasons for assuming that the adoption of the bipartite construction into Berber was primarily the result of borrowing, rather than imposition, are as follows. Historical evidence (see, e.g., Abun-Nasr 1987, Julien 1961) suggests that original ethnic Arabs have always been a small military-political elite in north Africa. Numbers grew somewhat in the eleventh century when conglomerations of Arab tribes such as the Banu Hilal migrated into the region, but these would have remained a minority.

This historical picture has been substantiated by a series of recent population genetic studies that have demonstrated that there is little or no significant genetic difference between those people in northwest Africa today who speak Berber and those who do not, or between those who identify themselves as culturally Berber and those who identify themselves as culturally Arab. At the same time, all of these groups differ significantly and in the same ways from today's inhabitants of the Middle East and Europe, including southern Spain (Bosch et al. 2000, Fadlaoui-Zid et al. 2004, Manni et al. 2002).

The conclusion we can draw from these facts is that the importance in this region of the Arabic language, Islam and other aspects of Arab culture is first and foremost the result of a cultural shift by autochthonous Berbers, rather than large-scale demographic replacement by, and intermarriage with, individuals of Arab genetic inheritance. As such, the numbers of genetically Arab L1 Arabic speakers will always have been small, and, given their status as an elite in a society whose cultural, religious and political institutions were being steadily arabized, the putative

acquisition of Berber as an L2 by a subset of these seems unlikely to have been the source of significant structural effects on the L1 Berber varieties of the masses.

However, ethnic origin is evidently not a sure predictor of linguistic dominance. It seems likely that, in the period when many Berbers started to shift to Arabic, at least some individuals were exposed to insufficient Berber during childhood to acquire it as an L1, but subsequently went on to learn it as an L2. Such individuals would then be expected to impose elements of their L1 Arabic onto their L2 Berber. However, to the extent that such individuals existed, they are likely to have been few in number and relatively uninfluential for the following reasons. Note that, although many peripheral varieties of Berber are in serious decline, there remain well over ten million native speakers of varieties with the stage II negative construction in Morocco and Algeria. A case could perhaps be made concerning some of the peripheral varieties, for example in Tunisia and Libya, that numbers of L1 speakers have been sufficiently small for sufficiently long that those who failed to acquire Berber as an L1 but did then learn it as an L2 could have represented a significant proportion of speakers of that variety and thus have made a non-trivial contribution to the PLD of children acquiring that variety as an L1. However, such speculations seem less plausible for the more widely-spoken varieties. Any such imperfect learners will always have been on the peripheries of these speech communities (usually having migrated from rural areas to predominately non-Berber-speaking urban centres; cf. Crawford 2002, Hoffman 2008) and until very recently the Berber language has been actively repressed rather than encouraged in the political, commercial and educational spheres (Hoffman 2006). Thus it seems unlikely that historically there have been large numbers of L1 Arabic speakers learning Berber as an L2 and making extensive

use of their L2 Berber, whereas it is certain that the L2 acquisition of Arabic by L1 speakers of Berber has been an increasingly widespread phenomenon in northwest Africa for almost a millennium.

Recall also the arguments given in section 3.4.1 against the idea that a bilingual speaker's dominance relations can be reversed under non-pathological conditions. If we accept these arguments then the many L1 Berber speakers who, over the years, have learnt Arabic as an L2 and perhaps started to make more use of Arabic than Berber in their daily lives will, even so, not have become dominant in Arabic, and could not therefore have exercised SL agentivity when speaking Berber. Arabic structures and lexicon will undoubtedly have often been more accessible to them when speaking Berber (thus making borrowing likely to occur; cf. chapter 4) but the character of their knowledge of their L1 Berber and their L2 Arabic will have remained distinct.

Hence it seems reasonable to conclude that the development of the bipartite negative construction in Berber was primarily the result of borrowing from L2 Arabic on the part of L1 Berber speakers. By contrast, in the following section I will develop the argument already mooted in chapter 2 that the prior development of this construction in Arabic was primarily the result of imposition on the part of L1 Coptic speakers.

A second likely example of grammatical borrowing comes from the Karnataka Saraswat Brahmin dialect of the Indo-Aryan language Konkani (henceforth KSKo), whose relativization strategy has apparently been remodelled on the pattern of the Dravidian language Kannada (Nadkarni 1975). KSKo features two types of clausal relativization strategy. One appears to be part of the common genetic inheritance of

the South-Western group of Indo-Aryan languages (including Marathi and Gujarati), although in KSKo this type is restricted to sentences where the relative pronoun is the subject of the relative clause (Nadkarni 1975: 678). This type, with its Marathi counterpart, is illustrated in (78):

- (78) (a) Marathi: jo mhātāra pepar vācət āhe to ḍāktār āhe
 (b) KSKo: jo mhāntāro pepar vāccat āssa to ḍāktāru āssa
 REL old.man paper reading is that doctor is
 ‘The old man who is reading a newspaper is a doctor.’
 (Nadkarni 1975: 675)

The major type of clausal relativization strategy in KSKo, in which the relative pronoun can be in any case and not just the subject of the relative clause, is strikingly different to that shown in (78), and is not found in the other Indo-Aryan languages. First of all, the relative pronoun used in this strategy is homophonous with the interrogative pronoun used elsewhere. Secondly we find an obligatory relative-clause-final suffix *-ki*.⁵⁵ This has close parallels with the Kannada relative clause, which is also introduced with a pronoun that is homophonous with the interrogative pronoun, and which features an obligatory clause-final suffix *-ō*, as shown in (79).

⁵⁵ This particle is optionally present in the KSKo Indo-Aryan type relativization strategy shown in (78). It is absent from the cognate constructions in Marathi and Gujarati (Nadkarni 1975: 675).

(79) (a) Kannada: yāva mudukanu pēpar ōdutta iddān-ō

(b) KSKo: khanco mhāntāro pepar vāccat āssa-ki

which old.man paper reading is-PRT

Kannada: avanu ḍākṭaranu iddāne

KSKo: to ḍākṭaru āssa

that doctor is

‘The old man who is reading a newspaper is a doctor’

(Nadkarni 1975: 674)

Relative clause-final *-ki* and *-ō* also have parallel etymologies. Kannada *-ō* is also used as an interrogative main clause complementizer, that is, it marks yes-no questions:

(80) (a) rāmanu sāleyinda bandanu

Ram from school came

‘Ram came from school.’

(b) rāmanu sāleyinda bandan-ō

Ram from school came-Q

‘Did Ram come from school?’ (Nadkarni 1975: 676)

It seems likely that the use of Kannada *-ō* as a relative clause complementizer represents an extension of its use as a main clause interrogative complementizer, following the crosslinguistically common development of relative clause markers from

question words (Lehmann 1984). In the same way, KSKo *-ki* also functions as an interrogative main clause complementizer:

- (81) (a) *rāmu skulāthāvnu āylo*
 Ram from.school came
 'Ram came from school.'
- (b) *rāmu skulāthāvnu āylo-ki*
 Ram from.school came-Q
 'Did Ram come from school?' (Nadkarni 1975: 676)

Other Indo-Aryan languages do not feature *-ki* in either of these functions, but “*-ki/ke* occurs in all [Indo-Aryan] languages as a[n embedded declarative clause] complementizer” (Nadkarni 1975: 676). Whether KSKo first extended the use of this embedded declarative clause complementizer to relative clause or interrogative contexts or perhaps both simultaneously is unclear, but it seems likely, given the very close patterning with Kannada *-ō*, which is not found in other Indo-Aryan languages, that it was due to transfer from Kannada that this extension took place (though, again, there seems to be no evidence in such cases that would count as actual proof of transfer rather than purely internal innovation).

What we know of the sociolinguistic situation of KSKo again makes it more plausible that this is a case of borrowing rather than imposition. Firstly, although KSKo speakers have for 400 hundred years been bilingual in Kannada, which is the official language in Karnataka (Nadkarni 1975: 673), “there has never existed a sizable group of native speakers of Kannada who also speak Konkani” (Nadkarni

1975: 679). Thus it is rather implausible that the source of any transfer from Kannada into Konkani should have been outside the Karnataka Saraswat Brahmin community.

It is also doubtful that the change in the KSKo relative clause could have come about as a result of imposition by members of the KSKo community who failed to acquire L1 competence in their community's language. One would expect such a situation to arise in the case of a moribund language which enjoys little social prestige – that is, a language whose speakers are shifting away from it, as we have seen is the case with certain Berber varieties. However, this is not the case with KSKo, despite its small number of speakers: "... socially and economically, Saraswat Brahmins have generally been in a prestigious position in Karnataka... Their orientation to Kannada has always been purely instrumental: they learned Kannada because they happened to live in a Kannada-speaking region" (Nadkarni 1975: 680). Thus a scenario whereby the innovation in the KSKo relative clause is due to borrowing under RL agentivity from Kannada looks to be the most plausible explanation in view of the available data.

This and the case of stage II negation in Berber illustrate cases in which the combination of linguistic and sociolinguistic data available make it more plausible to assume that the transfer of abstract syntactic structure has taken place under RL agentivity than under SL agentivity. This, then, is the first part of the argument against the rigid view advocated by Van Coetsem (1988, 2000) and Winford (2005), on which grammar (i.e. syntax and morphology) is claimed, as a whole, to be more susceptible to imposition than to borrowing. The aim of the following subsection is to argue for a more nuanced approach to syntactic imposition, which acknowledges that different syntactic structures exhibit different degrees of susceptibility to imposition,

and which offers an explanation of these differences based on the approach to L2 acquisition outlined in section 3.4.1.

3.4.2.2 Grammatical imposition

It was argued in section 3.4.1 that ‘the logical problem of L2 acquisition’ requires only the availability of adults’ general problem-solving and analytical abilities along with knowledge of their L1s. But what exactly is the role that knowledge of an L1 plays in L2 acquisition? There have been numerous suggestions on this issue put forward in the literature, of which one of the more popular is the Full Transfer/Full Access hypothesis of Schwartz and Sprouse (1994, 1996). This model, in contrast to Bley-Vroman’s Fundamental Difference hypothesis defended in section 3.4.1, maintains that UG is as central to L2 acquisition as it is to L1 acquisition, with the only difference being that the initial state of the language acquisition device when it comes to the task of L2 acquisition is that parameters are already set in accordance with the L1 grammar. The greater effort typically required for L2 acquisition is then associated with the process of resetting the relevant parameters, and the degree of imposition that L2 acquirers display is thus in inverse proportion to their progress in learning how the grammar of the L2 differs from that of their L1. How this model accounts for the apparent indeterminacy in end-state L2 grammars and the correlation between native-like ultimate attainment and high levels of verbal analytical ability is unclear, however. Moreover, it appears, at least in its simplest form, to make inaccurate predictions concerning the prevalence of imposition and the presence of structures in interlanguage that are not in accordance with the grammar of either the L1 or the L2.

This can be seen on the basis of a number of studies of adult L2 acquisition of basic elements of word order, such as the placement of verbs in the clause, and negation with respect to the verb. For example, Stauble's (1984) comparative study of Japanese and Spanish learners of English, mentioned in section 2.4, found that they exhibited similar patterns of development, with both sets of learners at early stages of acquisition expressing negation by placing *no/not* before the verb, despite the fact that negation follows the verb in Japanese, while it precedes it Spanish. If Japanese learners were imposing the "stable" grammar of their L1 on the RL English, one would expect them to preserve the Japanese VNeg order – however they appear not to.

Further evidence comes from studies on the acquisition of clausal word order. In the acquisition of a language such as German, which has underlying OV order combined with a verb second (V2) rule in main clauses, speakers whose L1 has SVO order (as in Romance languages) tend to assume that German has SVO order too, and regularly use this word order in subordinate clauses, where standard German would have SOV (Clahsen and Muysken 1986; Möhring and Meisel 2003). This fact would be consistent with either an imposition story, or a story which says that L2 learners simply analyse the surface strings of their L2 in a linear fashion (see below), since the V2 rule provides learners using the latter strategy with evidence that German has SVO order. However, the imposition story is undermined by the fact L1 Turkish speakers learning German and Dutch (which, like German has OV underlyingly and V2 in main clauses) also regularly use the SVO pattern in subordinate clauses, as in (82), despite the fact that this structure is found neither in

their L2 input nor in their L1, Turkish being consistently OV (Clahsen and Muysken 1986: 104):

(82) Wenn sie will gehen...

when she wants go.INF

'When she wants to go...' (Clahsen and Muysken 1986: 109, their (12))

This phenomenon, whereby L2 acquirers produce interlanguage with target-deviant word order when simply transferring the word order of their L1 would have resulted in target-like interlanguage, can also be seen in the L2 acquisition of V2 languages by individuals whose L1s also obey the V2 rule. For example, Håkansson, Pienemann and Sayehli (2002) and Pienemann and Håkansson (2007) show that L1 Swedish speakers learning German as an L2 routinely fail to apply the V2 rule in their interlanguage German, despite the fact that Swedish is also a V2 language. If they had simply imposed the V2 rule from their L1 onto their L2 then they would have produced target-like word order in their interlanguage. As it is, they at least go through a stage of producing interlanguage that is ungrammatical in both their L1 and the L2.

These data suggest that an alternative model of L2 acquisition is required, in which the wholesale imposition of the syntax of the L1 onto the L2 is not predicted. Several caveats are required here, however. Firstly, these arguments should not be interpreted as generalizing to all elements of abstract syntactic structure: there is no doubt that certain elements of word order, such as adverb placement for example, are very commonly imposed (e.g., White 1991). The interesting point, which we

return to shortly, is that the same should not be true, to the same extent, of more fundamental aspects of word order such as the position of the verb in the clause.

Secondly, it should be said that even with respect to basic word order the available data are conflicting. A clear example comes from Schwartz (1998), who reports on a study by Haznedar (1995) in which a child with Turkish L1 produces OV structures in his L2 English. Similarly, Odlin (1989) reports on a study by Nagara (1972) in which the Pidgin English of L1 Japanese speakers in Hawai'i is also shown to contain some instances of OV structures (the basic word order in Japanese). These and other examples in the literature of imposition of basic word order share an important property, however. In contrast to the imposition of adverb order,⁵⁶ for example, which appears to be a feature of even near-native interlanguage, the imposition of basic word order appears only to be a feature of the earliest stages of L2 acquisition. In Haznedar's study the child in question regularly produces English sentences with OV order during the first two months of his exposure to English only. Thereafter basic word order is almost entirely target-like (Haznedar 1995: table 1; Schwartz 1998: 137). In the same way, Siegel (2003), also commenting on the study by Nagara (1972), notes that L2 acquisition of basic word order tends to be very rapid, and points out that, to the extent that L1 Japanese speakers did produce OV structures in the early days of their exposure to Hawai'i Pidgin English, this "was a

⁵⁶ This phenomenon is common, for example, in the L2 German speech of L1 English speakers, where one might commonly find ungrammatical strings such as (i) instead of the grammatical (ii):

- (i) *Felix lief nach Hause glücklich heute morgen.
- (ii) Felix lief heute morgen glücklich nach Hause.
Felix walked home happily this morning.

transient feature and was not found in the stabilized Hawai'i Pidgin English" (Siegel 2003: 195 fn.10). At the same time Odlin (1989: 94) notes that there seems to be "an inverse relation between transfer [of basic word order] and proficiency."

Finally, note that a clear distinction needs to be made between abstract syntactic structure such as basic word order (which seems not to be particularly susceptible to imposition) on the one hand, and the argument structure, referential or other properties associated with specific lexical items on the other, which I take, with Winford (2005: 380), to be quite readily imposable. There is certainly a syntactic component to such cases of imposition, but the important point is that here it has a clear lexical basis. We can illustrate this first by considering the examples (70) and (71) above. In (70), the SL-dominant speaker associates an item in the RL, *emprunter*, with an already existing lemma in her mental lexicon, *borrow*, and therefore endows this RL item with all the syntactic and semantic properties of its perceived SL equivalent until such time as sufficient exposure to the RL causes her to make the necessary alterations to her representation of the syntax and semantics of this item. In (71) we have another instance of lexically-based imposition, this time involving what Heine and Kuteva (2003, 2005) have called 'polysemy copying'. The sound string *uns* in German is polysemous, serving as the phonetic manifestation both of the accusative/dative first person plural pronoun (i.e., 'us' in English), and a reciprocal bound anaphor co-indexed with a preceding first person plural subject (i.e., one of the functions of 'each other' in English). The SL German-dominant speaker in (71), therefore, is exposed to the English sound string *us* and correctly establishes that it functions as the accusative/dative first person plural pronoun. However, she also wrongly assumes that it is polysemous in the same way as the German sound

string *uns*, and can therefore be used to express a reciprocal bound anaphor co-indexed with a preceding first person plural subject.

I now want to argue that another instance of this same phenomenon of lexically-based syntactic imposition is the transfer of bipartite negation from Coptic into Arabic. As with the Berber and KSKo developments discussed in the previous section, the case needs to be made here that it was indeed primarily imposition and not borrowing that was responsible for this development. Given the current paucity of information about the precise sociolinguistic situation in Egypt in the period in question, however, as in the Berber case, we are necessarily operating here with reasonable assumptions rather than proven facts. What information we do have, though, makes it reasonable to assume (as we did implicitly in the discussion surrounding Figure 2.2) that this change occurred under SL agentivity.

Firstly, it is clear that the long-term trend was for Coptic speakers to shift to Arabic and not vice versa, since Coptic eventually died out as a spoken language. Moreover, this shift appears to have started early, at least in official discourse, as Arabic was substituted for Coptic in all state affairs already in 706 (Solihin 1991: 12) – though the Arabic used in official matters will have been an early form of Standard Arabic of course.

Secondly, as in the rest of north Africa, indications are that the Arabs remained a small military-political elite for some considerable time in Egypt, and even if some of them did learn Coptic as an L2, these would have represented a fraction of the numbers of Copts who learnt Arabic as an L2.

Lastly, we noted in chapter 2 that the Arabic stage II negative construction cannot be considered a calque on that of Coptic, and structurally it is rather different. Even if

there were L1 Arabic speakers who learnt Coptic as an L2 and were thereby somehow prompted to replicate the Coptic stage II construction in their L1 Arabic, how or why they should have recruited their indefinite pronoun for this purpose and made it immediately postverbal, rather than, say, recruiting an existing Arabic negator or borrowing Coptic *an* and keeping it in its predominately clause-late position, is unclear. On the other hand, the motivation for L2 speakers of Arabic (with L1 Coptic) having done this is reasonably straightforward, given the above considerations.

First of all, the position of Arabic *mā* immediately preceding the verbal group (rarely some other focus of negation) is entirely expected from the point of view of a learner familiar with Coptic: it directly parallels the position of *ən*. Thus far imposition and successful acquisition are indistinguishable. Having identified a straightforward translation equivalent of *ən*, these learners will then, all else being equal, expect to find a second, postverbal item that is the equivalent of *ən*. By hypothesis, the item which they interpreted as being the equivalent of *ən* was *šay?*, which, as we have seen, occurred predominantly in the scope of negation in the Classical Arabic period. How this imposition then led to a change in the grammars of L1 Arabic speakers is spelt out in more detail in chapter 4. The crucial point for present purposes is that this scenario sees the imposition of stage II negation on Arabic as a case of lexical transfer as with the examples in (70) and (71). The syntactic and semantic properties of an item in the SL (*ən*) are transferred to an item in the RL (*šay?*), as is typical in L2 acquisition. Note that this sharply distinguishes this scenario from the superficially similar scenario investigated by Stauble (1984), whereby L1 Japanese speakers fail to impose VNeg order on their L2 English (which of course has (Aux)NegV order).

Imposition of VNeg order on NegV English would have involved these speakers' substantially ignoring a rather salient piece of word order. By contrast, in the scenario for the transfer of stage II negation to Arabic proposed here, there is no imposition of word order, merely the maintenance of the preverbal negator in its original position and the identification with the postverbal negator of a postverbal item that is present in the input.

Taken together, then, the preceding discussion suggests that the true picture is more complex than a statement such as "imposition tends to affect the RL grammar" would suggest: the imposition of syntactic properties of lexical items as well as various elements of abstract syntactic structure is certainly widespread, while other elements, such as basic word order, seem to be less readily imposed.

We can make sense of these differences in terms of the approach to L2 acquisition defended in section 3.4.1. If speakers draw on general problem-solving abilities in acquiring their L2, and thus, "rather than using structure-dependent operations constrained by UG [or, *a fortiori*, by the parameter settings of their L1 – CL], resort to linear sequencing strategies which apply to surface strings" (Meisel 1997: 258; cf. also Bley-Vroman 2009 on the importance of "shallow parses" for L2 comprehension and acquisition), then the empirically attested patterns of imposition are to be expected. Many studies of L2 acquisition have found that basic word order is not commonly imposed, and, where it is, that this is typically only for a brief period in the earliest stages of acquisition. This is presumably because virtually every utterance of the L2 that the acquirer is exposed to will contain evidence of the basic word order in that language, whereas this is not true for verbal adjuncts, for instance, and especially not for how they are ordered relative to one another (cf. footnote 56).

As such, the basic word order of an L2 is likely to be more salient to the acquirer than other aspects of syntax (cf. Siegel 2003), for which the evidence is also there in the input (just as it is for L1 acquirers) but is obscure to all L2 acquirers except those with high levels of verbal analytical ability.

However, clearly nothing about an L2 is salient to the acquirer when she has had little or no exposure to it. It seems natural to suppose that, in the absence of (sufficiently salient) evidence for some feature of the syntax of the L2, the acquirer will, consciously or unconsciously, draw on her only other source of knowledge of language and hypothesize that this feature is identical to its counterpart in her L1. This would explain why some of the most salient aspects of abstract syntactic structure appear only to be imposed for a brief period at the beginning of L2 acquisition, if at all. At the same time we have seen that more complex word order rules, such as the basic OV order combined with V2 in main clauses as found in German pose greater problems for L2 acquirers. Whether the word order rules of these acquirers' L1s match that of the L2 partially, entirely, or not at all, the result after the earliest stages of L2 acquisition appears to be the same: V2 in main clauses is interpreted as SVO order, which is then extended to embedded clauses. This could potentially be interpreted as imposition on the part of speakers whose L1s are rigidly SVO, but with speakers whose L1s are OV (or OV + V2) something else is at work. I propose the label 'restructuring' for changes which a speaker makes to an L2 that are the result not of imposition but of interpreting the L2 input in a way that a child acquiring an L1 by means of UG would not. The possibility of this process having been operative in the well-known change from OV to VO in English is discussed in section 3.4.3.

In summary, then, the empirical evidence currently available suggests that the idea of the stability gradient, and especially Winford's (2005: 377) claim that "borrowing [...] does not usually affect the RL grammar, while imposition, on the other hand, tends to do so", is too simplistic and could be misleading as it stands. It does not seem justified to assume that a historical change in basic word order, for example, is more likely to have been the result of imposition than of borrowing. If anything, the reverse may be true: for the language of an entire speech community to undergo contact-induced syntactic change the contact must presumably be intensive and sustained. However, these are exactly the conditions under which imposition of basic word order appears, on the empirical evidence currently available, not, or hardly, to occur. In fact, the sensible way to proceed in such cases must be to make the available sociolinguistic information about the contact situation criterial in establishing whether a change was due to borrowing, imposition, or perhaps both.

To the extent that different domains of language are characteristically associated more with one of the two main transfer types than the other, this seems likely to have its own independent cause for each given domain, but can in part be explained in terms of the very different psychological processes underlying L1 attrition and L2 acquisition. The prevalence of phonological imposition, for example, is likely to have a physiological or motor neural basis. Van Coetsem (1988: 27-8) in fact acknowledges this, citing Walsh and Diller's (1981: 18) explanation (also cited in Köpcke 2004: 10): "Low-order processes such as pronunciation are dependent on the early maturing and less adaptive macroneural circuits, which makes foreign accent difficult to overcome after childhood." The pervasiveness of lexical borrowing, by contrast, is likely to have much more to do with processes of cultural exchange and

the borrowability of objects and concepts along with their linguistic labels than with properties of the mental lexicon. Grammar (syntax and morphology), taken as a whole, cannot simply be claimed to be particularly prone to imposition, and this seems to be because L2 acquirers use general cognitive mechanisms to analyse the surface strings of the L2, rather than simply assuming it has the same hierarchical structure as their L1. The argument structure and referential properties of lexical items, on the other hand, do seem to be quite readily imposed, presumably because of the identification of RL items with lemmata in the mental lexicon of the SL-dominant speaker. Finally, there does seem to be plausible evidence for the existence of grammatical borrowing, though it is not yet clear whether there are any differences in the prevalence of borrowing between, for example, basic word order and adverb order, as we have seen for imposition. How and why syntactic borrowing occurs forms the core of the discussion in section 4.1.

3.4.3 Restructuring

We have already seen the answer to the question of whether borrowing and imposition exhaust the possible types of contact-induced change: they do not. In the previous section I gave the label 'restructuring' to changes which a speaker makes to an L2 that cannot be seen as the transfer of patterns or material from their L1. Such changes are clearly the result of contact, but are perhaps better described as occurring under L2-agentivity rather than SL-agentivity since they do not involve transfer from any SL. Nevertheless, there is a clear link between changes due to restructuring and changes due to imposition in that agents of change in both cases are L2 acquirers of the RL and are thus not dominant in it.

As we have seen, there is no doubt that second language learners produce structures which are target-deviant but which do not reflect the grammars of their L1s. As an illustration of how this process could be an input to change at the level of the speech community let us take the famous but controversial example of the change from OV to VO in English.

Weerman (1993) makes the case for the well-known change from basic OV to VO order in English being contact-induced. Specifically, he draws a number of parallels between the synchronic syntax of Middle Dutch and Old English (the most prominent being underlying OV order, V2 in main clauses and the possibility of extraposition) and asks why, given the similarity between these two systems, one should have become verb-medial while the other remained verb-final. (Although Weerman makes no specific reference to it, this is, once again, an instance of Weinreich et al.'s (1968) actuation problem). His answer is that the crucial difference was that English was influenced by considerable numbers of native speakers of Old Norse who learnt Old English as a second language during the Viking occupation of northern and eastern England from the ninth to the eleventh centuries, whereas Dutch had no comparable influence. More recently, the idea that the change to VO order in English was the result of Norse influence has been made the subject of a book-length study by Trips (2002).⁵⁷ On the other hand there have been various purely internal accounts of this change, most notably Lightfoot (1991) and Kiparsky (1996), while Roberts (2007: 391-399) specifically rejects the idea of a role for contact in this change. Roberts'

⁵⁷ On a closely related topic, Kroch and Taylor (1997) argue that the lack of V2 in many types of embedded clause in northern Middle English (as compared to the presence of V2 in those same clause types in southern Middle English) is also due to contact with Norse.

critique is based on two main arguments which are worth examining here because, in the light of some of the points already made, they do not seem to be valid.

Roberts' (2007: 393) first point is that, for the contact story to work "we must assume that [Old Norse] was a VO language." He then cites various sources to demonstrate that this was in fact probably not the case, rather that Old Norse was most likely either OV or mixed OV/VO. However, as we have seen, and as Weerman (1993: 924) points out, citing Clahsen and Muysken (1986), this assumption is unjustified. Let us take it that Roberts is right that Old Norse was essentially OV. What Clahsen and Muysken (1986) demonstrate is that L1 speakers of OV languages learning as an L2 a language that is OV with V2 only in main clauses tend to overgeneralize VO order to embedded clauses in just the same way as second language learners whose L1s are VO. So it seems that whether Old Norse was OV or VO during this period is in fact has little bearing on whether or not contact with Norse was a cause of the shift in English verb placement.

Roberts' (2007: 395) second criticism of the contact story is that the West Saxon variety of Old English already shows evidence of mixed OV/VO order in the late ninth century; that is, at a time before this can plausibly be attributed to Norse influence. There are two points to be made in response to this. First of all, as has been argued by Thomason (2007) among others, the fact that there are internal factors which are (hypothesized to be) sufficient to cause a given change in a given language does not, by itself, imply that external factors cannot also have contributed to this change. Thus, the fact that Old English might, in the hypothetical absence of contact, conceivably have developed its current rigid VO order through successive reanalyses

of earlier mixed OV/VO orders does not mean that the extensive contact that did take place did not increase this potential and help to make it a reality.

The second point concerns the core of Weerman's (1993) argument in favour of the contact story. This is that none of the syntactic properties suggested, for example, by Lightfoot (1991) as causes for reanalysis of English as VO (e.g. separable verbal prefixes appearing to the left of the object or in second position and negation being a verbal prefix, in addition to V2 and extraposition) are in fact sufficient in themselves to cause such a reanalysis. We can say this with confidence because Middle Dutch shared all of these properties and yet Dutch never became VO. This same argument applies to Roberts' second point. The fact that early West Saxon displays VO orders in embedded clauses (83) is inconclusive because we find the same thing in Middle Dutch (84), which never became VO.⁵⁸

(83) *pæt he andette his scrite ealle his synna*
 that he confess.PST.3SG.SBJV his confessor all his sins
 'that he confess to his confessor all his sins.'

(9th-century West Saxon; Weerman's (15b))

(84) *dat si ontmoetten ene ioncfrouwe*
 that they meet.3PL.PST a lady
 'that they met a lady.'

(Middle Dutch; Weerman's (15c))

⁵⁸ In fact the same argument would appear to apply to Kiparsky's (1996) internal account of the shift to VO order in English. This account claims that the change took place only in those Germanic languages which allowed verb-fronting in embedded clauses. Why Middle Dutch, which also appears to have allowed this operation, never became VO is unclear on Kiparsky's account (which does not mention Middle Dutch at all).

Thus it seems at least as reasonable to adopt Weerman's (1993) proposal that the change from OV to VO order in English involved contact with non-native speakers, as it does to insist that it was a purely internal change, particularly if one takes Longobardi's (2001: 278) version of the Inertia Principle seriously: "syntactic change should not arise, unless it can be shown to be *caused*" (emphasis in original).

If this is accepted then we have a clear example here of restructuring: changes to an RL made by L2 acquirers which do not involve transfer from an SL. This, then, is the third major type of contact-induced grammatical change, alongside borrowing (under RL agentivity) and imposition (under SL agentivity).

Having introduced and exemplified these three major types of contact-induced change in the present chapter, chapter 4 spells out in more detail how it was that contact could have led to change in the mental grammars of L1 speakers in the cases examined thus far, focussing in particular on the cases of stage II negation in Arabic and Berber.

4 Mechanisms of contact-induced change

It was argued in the previous chapter that transfer of lexicon, phonology and syntax are likely each to have their own independent causes, of which the appearance of a stability gradient is an epiphenomenon. It was also argued that different aspects of syntax exhibit different levels of susceptibility to imposition, without obviously showing an inverse level of susceptibility to borrowing, as the idea of the stability gradient would lead one to expect. This points to the conclusion that there might not be any particularly illuminating way in which the different domains of syntax, phonology and the lexicon can be brought together in a single unified framework for understanding contact-induced change. The purpose of this chapter is more modest: to explore in more detail some of the mechanisms potentially underlying the syntactic changes already described for Arabic and Berber, as well as for Karnataka Saraswat Konkani (KSKo) and English, and to show how these mechanisms can be integrated into the standard innatist model of abductive change (after Andersen 1973).

In essence, this model states that children acquiring their L1 are the locus of change, expressed in generative work since the 1980s as change in parameter settings. This comes about when the set of primary linguistic data (PLD), or the trigger experience, to which the child acquiring the language is exposed differs with respect to that on the basis of which older groups of speakers acquired their own grammars. This then causes the child to abduce a different grammar (different parameter settings) to that of older groups. This alteration in PLD could have various extra-syntactic causes, such as phonological change, change in frequency of use of various structures, and, most importantly for our purposes, language contact. Put

another way, then, the aim of the present chapter is to try to make more explicit some of the ways in which contact can bring about changes in PLD, leading to change in grammars through L1 acquisition.

It is worth noting that the outlines of a similar proposal have recently been given by Roberts (2007: 236-7, 390-1). There are important differences, however, between that proposal and the one given here.

Roberts distinguishes two types of contact-induced change in PLD: 'direct' and 'indirect'. He does not discuss borrowing in Van Coetsem's sense. For Roberts, direct contact appears to be where tokens from separate languages can be combined into a single set of PLD for a child acquiring a single language. However, given the now widely-acknowledged fact that children in a bilingual environment achieve early and successful differentiation of the languages they are exposed to (cf. footnote 54), this type of contact-induced change in PLD would appear to be relevant only to contact between closely related dialects and not for mutually unintelligible languages.⁵⁹

Roberts' indirect contact, on the other hand, is equivalent to our SL/L2 agentivity: the interlanguage of L2 acquirers of a given RL will typically be markedly different from the performance of L1 speakers of that RL (including features transferred from the SL, these speakers' L1) and this interlanguage forms (part of) the PLD for children acquiring the RL as an L1.

We have already seen some of the arguments for why interlanguage is typically so different from L1 performance. Section 4.1 expands on these to give a more

⁵⁹ Quite what cues children use to distinguish different languages in their environment is beyond the scope of the present discussion, but phonological differences appear to play an important role from a very early age (Bosch and Sebastián-Gallés 2001).

explicit account of the development of stage II negation in Arabic, while section 4.2 focuses on the mechanisms potentially underlying borrowing in Berber and KSKo, especially L1 performance attrition and Heine and Kuteva's (2003, 2005) notion of 'contact-induced grammaticalization'. Section 4.3 then revisits the issue of L2 agentivity and restructuring in the discussion of the change in word order in English.

4.1 Coptic to Arabic imposition of stage II negation

We have seen in the previous discussion that second language learners will come to an L2 expecting to find in it features they are familiar with from their L1. If, however, they find no evidence for a particularly salient or common feature, they are unlikely to impose it. If, on the other hand, they do find what they take to be evidence for this feature, in the absence of correction they will be likely to impose it on their L2. Of course, on a view which sees L1 acquisition as the task of a dedicated mental module, while L2 acquisition uses more general problem-solving strategies, it is to be expected that what counts as evidence for a feature to L1 acquirers is quite different from what counts as evidence for a feature to L2 acquirers.

I propose that L1 speakers of Coptic learning Arabic as an L2 did indeed find evidence in the Arabic they were exposed to for what to them would have seemed a 'natural', stage II negative construction.⁶⁰ Typical sentences such as (85)b, for instance, might provide such evidence.

⁶⁰ In the absence of detailed information to the contrary, I assume for the time being that the processes that led to Jespersen's Cycle in the Arabic of Yemen and Oman were essentially the same as those described here for Egyptian Arabic, albeit, of course, with Modern South Arabian as the contact language rather than Coptic.

(85) (a) tašrab qahwa
 drink.IMPF.2MSG coffee

‘Would you like some coffee?’

(b) lā mā ašrab šay? qabl an-nawm
 no NEG drink.IMPF.1SG anything/NEG before the-sleep

Conservative interpretation: ‘No, I don’t drink anything before bed.’

Innovative interpretation: ‘No, I don’t drink coffee before bed.’

The conservative interpretation of (85)b is that the question is about coffee, whereas the answer is about anything, including coffee. If, however, an L2 acquirer were expecting to find negation expressed by two elements, one either side of the verb, it seems plausible that they could interpret *šay?* here as a postverbal negator and assume that (85)b features a null object (as would be common in Arabic and other languages when there is a salient referent in the discourse for such an object). On this interpretation both the question and the answer would concern coffee, rather than the answer concerning anything at all. Recall also that *šay?(an)* functioning as an adverb or an argument in Classical Arabic is predominantly found in the scope of negation, making the innovative interpretation of (85)b all the more plausible.

Clearly such contexts would have been available in all varieties of Arabic and, in principle, children acquiring Arabic as an L1 could have been equally susceptible to this reanalysis as adult L2 learners. However, as we have seen, only a subset of Arabic varieties actually underwent Jespersen’s Cycle. This must mean that contexts as in (85) have, in fact, always been transparent to L1 acquirers of stage I varieties: despite the apparent ambiguity, they have not found evidence to suggest that *šay?*

(or a reduced form *šī*, or some other synonymous item) is anything other than an indefinite pronoun (or adverb) here. It seems that only L2 acquirers of Arabic whose L1 features a stage II construction are liable to find such contexts ambiguous and decide that they support an analysis of Arabic such that it too features a stage II construction.

However, L2 learners are still bound to some extent by the (surface) syntax of the tokens of the L2 they are exposed to. As we have seen, it is apparently only in cases of radical underexposure to the L2, if at all, that speakers will simply relexify their L1 with its syntax completely intact (see also Lefebvre 2001). Ordinarily, there will be some syntactic features of the L2 that are sufficiently salient to the L2 acquirer such that they are not simply replaced by their equivalents from these speakers' L1s. So it appears that in the present case, while native speakers of Coptic found evidence in structures such as (85)b that Arabic, like Coptic, also had a stage II negative construction, at the same time they found evidence that in Arabic the postverbal element is routinely verb-adjacent, rather than occasionally so as in Coptic, so they did not simply transfer the syntax of their stage II construction wholesale into their L2 Arabic (cf. Figure 2.2).

If sufficient speakers then incorporated these analyses into their production of their L2 Arabic, the conditions are there for this new hybrid construction to make its way into the L1 grammars of Arabic speakers via child language acquisition. First of all, as use of Arabic became more and more common in the Coptic community, increasing numbers of children in this community will have acquired Arabic as an L1 simultaneously with Coptic. The PLD on the basis of which these children acquired Arabic as an L1 will have come partly from native speakers of Arabic, but an

important component will also have stemmed from the L2 production of native speakers of Coptic. Assuming these speakers' L2 Arabic speech contained the innovative stage II negative construction, this construction will also have formed part of the PLD for the children acquiring Arabic, so that they are liable to have acquired this construction as part of their native competence in Arabic, unlike the L1 acquirers of Arabic in other regions whose PLD did not contain a stage II negative construction.

At this stage we have a fully-fledged contact-induced change in the grammar of the L1 variety of Arabic spoken by younger members of the increasingly bilingual Coptic community. The mechanisms which I have argued underlie this and other changes due to imposition are summarized in Figure 4.1.

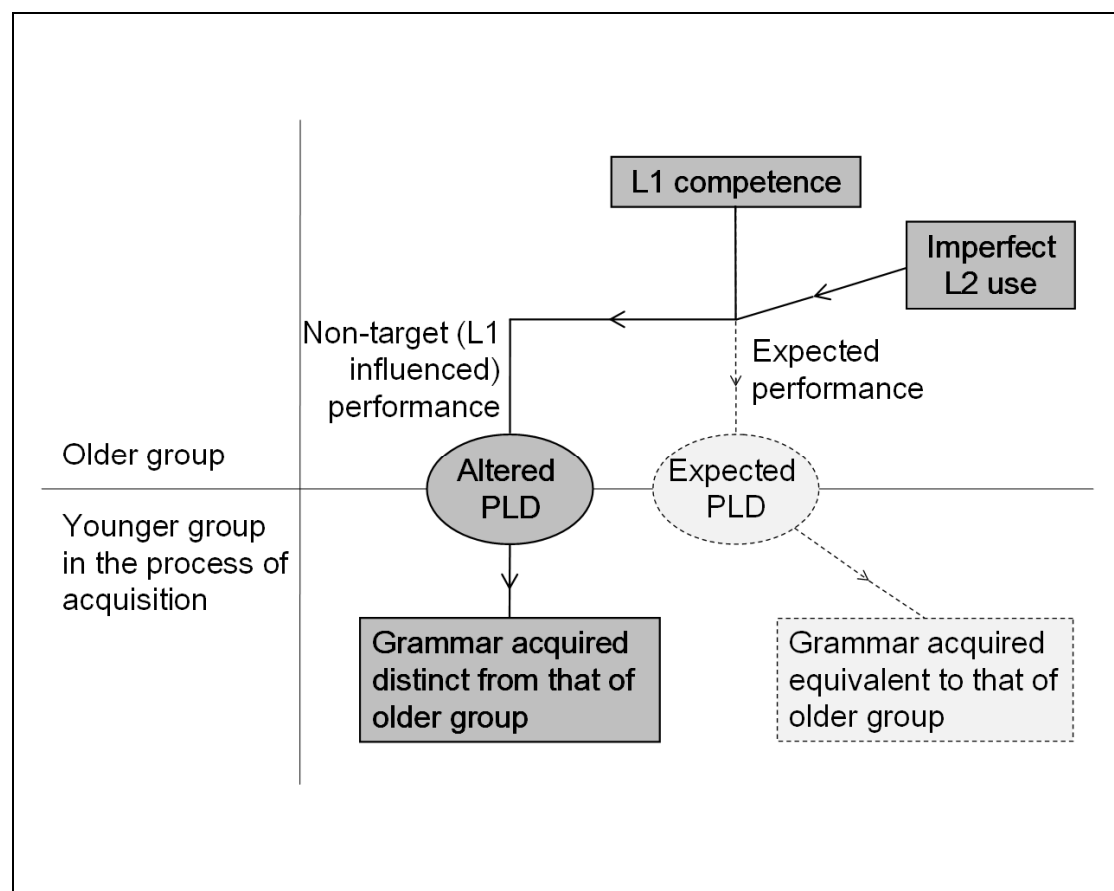


Figure 4.1

How imposition leads to change

From this point on, given the right sociolinguistic circumstances, the potential is there for the stage II negative construction to diffuse rapidly among mutually intelligible Arabic varieties. These circumstances are presumably that speakers of different varieties come into contact, and that there should be a sociolinguistic incentive for speakers of the recipient variety to adopt at least some features of the source variety. Since, before long, the numbers of Arabic speakers from the Coptic community will have greatly outnumbered the Arabs in Egypt, it is not surprising that the Coptic-influenced variety should have rapidly become the norm within Egypt as a whole. (Recall that there is a sharp divide in the Arabic-speaking world between the standard written language, which is highly conservative, and the spoken varieties, concerning which there tends to be little or no prescriptivism). From there, as mentioned in section 2.4.2, the diffusion westwards of stage II negation is also expected thanks to Cairo's influential position in north Africa as the seat of the Fatimid Empire and successive waves of migration from Arabia and the Levant via Egypt into western north Africa.

Thus we have the basis at least of a genuinely explanatory account of how contact in Egypt between Coptic, with a bipartite negative construction featuring a predominately clause-late second element, and Arabic, with its originally preverbal construction, resulted in the spread of a new Arabic bipartite construction, with a verb-adjacent second element, across the whole of coastal north Africa. The next section attempts a similarly explicit account of how, in turn, this construction spread from Arabic Berber.

4.2 Arabic to Berber borrowing of stage II negation

Even if there is no consensus in the second language acquisition literature as to the root causes of the divergence between an average second language learner's interlanguage and the speech of a native speaker of the same language, the divergence is nevertheless familiar and understood at an intuitive level at least. The same cannot be said for grammatical changes made to a language by native speakers of that language (changes under RL agentivity). The purpose of this section is to explore in some detail what makes these changes possible, and what motivates them, in order to be able to give an explicit account of the borrowing of stage II negation from Arabic on the part of Berber, as well as the borrowing of the Kannada relativization strategy in KSKo introduced in chapter 3.

4.2.1 L1 performance attrition

We saw in the earlier discussion of L1 attrition that, while there is plenty of evidence that bilingual speakers often show attrition in their L1 performance, theoretical considerations and a lack of empirical evidence thus far make it sensible to proceed on the assumption that L1 competence is not attrited in non-pathological bilingual individuals. What this means is that if we take syntactic change to be innovation in the (mental) grammars of individual speakers, then the L1 attrition exhibited by bilinguals is a trigger for innovation, in the form of perturbation to PLD, rather than the innovation itself. Without ruling out the possibility that there are other causes of transfer under RL agentivity (e.g., codeswitching perhaps; Backus 2005), L1 performance attrition would appear to be among the more important of these. It will therefore be instructive to consider here what it is that causes performance attrition,

before working through some examples of how it might act as an input to change in competence.

I suggest that L1 performance attrition is ultimately the result of an inherent psychological tendency to minimize the processing effort associated with the use of two distinct languages. This idea is by no means new. Nadkarni (1975: 381) expresses it thus:

“bilingualism is, after all, a psychological load – not so much because it requires knowing two language systems, but because, in a situation of intensive bilingualism, one is called upon to conduct communication through these distinct systems all the time, using now one system and now the other. In such a situation, the tendency toward lessening the psychological load is quite natural.”

In fact this idea is presaged in the work of Uriel Weinreich, albeit not explicitly in the cognitive terms we are aiming for here: “inasmuch as a language is a system of oppositions, a partial identification of those systems is to the bilingual a reduction of his linguistic burden” (Weinreich 1953: 8). A possible way in which such a reduction of a bilingual speaker’s cognitive (rather than “linguistic”) burden could be implemented is as follows.

Even if one accepts the “massively modular” view of cognition (Carruthers 2005), there is nevertheless clearly an important interaction between language and “habit”, or procedural memory, potentially expressible in terms of strength of neural connections. One might hypothesize a neural network associated with, say, a

morpheme or a syntactic structure (or a class of either of these) in a speaker's L2. If the strength of the connections between the units in this network is high due to repeated exposure to, and use of, the relevant morpheme/structure, one would expect there to be decreased processing costs associated with involving this same, already strongly activated network in L1 use too. Put simply, it pays to carry over habits formed in an L2 into one's L1. Anyone who has spent time immersed in an L2 environment will have anecdotal examples of the resulting urge (sometimes irresistible) to indulge L2 habits during L1 speech.

A similar suggestion is made by Köpke (2007), with a focus on speakers who make considerably more use of an L2 than their L1: where a speaker makes little use of and has little exposure to her L1 over an extended period, this will result in decreased activation of the neural networks associated with communication in the L1, manifesting itself to the consciousness of the speaker as a decrease in the accessibility of that L1. At the same time, the neural networks associated with communication in an L2 that is used frequently will be strongly activated, and therefore that L2 will be readily accessible. However, if in this situation the speaker in question wishes to express herself in the deactivated L1, this must be reactivated, and the strongly activated L2 must be inhibited. This simultaneous reactivation and inhibition will put pressure on the cognitive mechanisms controlling speech production, such as working memory. Where this pressure is moderate, normal L1 production will result. Where it is excessive, however, one would expect the full reactivation of the L1 and inhibition of the L2 to be unsuccessful and for L2-influenced performance attrition of the L1 to result. Bearing these ideas in mind, let

us now revisit the cases of apparent syntactic borrowing in Berber and KSKo described above.

4.2.2 Performance attrition in Berber and KSKo

In earlier work (Lucas 2007: 407-8) I suggested that that the transfer of stage II negation from Arabic to Berber can be accounted for in terms of Heine and Kuteva's (2003, 2005) notion of contact-induced grammaticalization, but without going into detail on this point. Let us now see whether this proposal can be made to work, and, if so, how.

Heine and Kuteva present the idea of contact-induced grammaticalization as an extension of grammaticalization theory from cases of internal, to cases of external grammatical change (2005: 1). They (2003: 533, 539) identify two varieties of contact-induced grammaticalization, which they call 'ordinary' contact-induced grammaticalization (86), and 'replica grammaticalization' (87), respectively:

(86) Ordinary contact-induced grammaticalization

- a. Speakers notice that in language *M* there is a grammatical category *Mx*.
- b. They create an equivalent category *Rx* in language *R* on the basis of the use patterns available in *R*.
- c. To this end, they draw on universal strategies of grammaticalization, using construction *Ry* in order to develop *Rx*.
- d. They grammaticalize *Ry* to *Rx*.

(87) Replica grammaticalization

- a. Speakers notice that in language *M* there is a grammatical category *Mx*.
- b. They create an equivalent category *Rx* in language *R*, using material available in *R*.
- c. To this end, they replicate a grammaticalization process they assume to have taken place in language *M*, using an analogical formula of the kind [*My* > *Mx*] : [*Ry* > *Rx*].
- d. They grammaticalize *Ry* to *Rx*.

Heine and Kuteva use the terms M(odell) language and R(eplica) language equivalently to our SL and RL respectively, so I will retain the latter terminology in the following discussion, except where directly quoting Heine and Kuteva. Leaving aside for the moment the immediately apparent problem of how we are to understand “notice” in this context, the process outlined in (87) appears to capture the case of Berber negation fairly well. As we have seen, Arabic is the SL and Berber the RL, and Berber speakers appear to have replicated a grammaticalization process they assume to have taken place in Arabic (word for ‘(any)thing’ > postverbal negator) in order to create a new bipartite negative construction. The fact that the postverbal negator was grammaticalized along similar paths in both Arabic and Berber makes this case fit (87) more closely than (86).

The fit is not perfect, however. Note that in both (86) and (87) Heine and Kuteva refer to the creation of grammatical *categories*. Given that they nowhere explicitly define what they mean by the use of the word ‘category’ in this context, one can only assume that it is in the standard sense, roughly: ‘abstract grammatical class or

semantic notion whose lexical realization is a set of items (free words or bound morphemes) which belong to this class in virtue of their shared syntactic properties'. This does seem to cover the majority of Heine and Kuteva's uses of this term, but, this being the case, it is surprising that they should primarily limit their attention to the *creation of new* grammatical categories. Certainly there has been no new category created in the present case, since, whatever its syntactic manifestation, Berber will always have had the category of negation – this I take to be a linguistic universal. But if this is correct, how are we to understand the following statement from Heine and Kuteva (2005: 121-2): "there is little evidence for some [...] categories arising as a result of language contact, such categories relating to personal deixis or negation." If Heine and Kuteva are using *categories* in the standard sense here, and it is true that having the category of negation is a linguistic universal, then it is hard to understand what could be meant by this statement. For the category of negation to arise in a language due to contact it would presumably have not to have existed in that language previously, which seems implausible. If, on the other hand, Heine and Kuteva are using *category* here and in (86)-(87) more in the sense of 'construction', then they appear to be wrong in saying that there is no evidence for negative constructions arising in this way, as the Coptic-Arabic and Berber-Arabic cases illustrate.

Moreover, returning to the problem of "notice" in (86)-(87), if we take this literally it would suggest that the processes described in (86)-(87) are necessarily conscious and intentional, with speakers engaged in these processes having a great deal of metalinguistic knowledge at their disposal (see Karatsareas 2007: 36, Gast and van

der Auwera 2009 for similar concerns). This hardly seems plausible, though it is possible that there is a degree of consciousness involved in these types of changes.

These reservations aside, Heine and Kuteva do seem to have identified a major type of syntactic borrowing (see Heine and Kuteva 2005 for numerous examples from a wide range of languages), their characterization of which we can adapt and integrate with the ideas developed here as follows. Because the case of Berber negation does not fit (86) and (87) in that it does not involve the creation of any new grammatical category, we need a third type of contact-induced grammaticalization which will accurately capture the Berber facts:

(88) Category-preserving contact-induced grammaticalization

- a. An RL expresses a grammatical category φ by means of construction Rx .
- b. An SL also spoken by RL-dominant speakers expresses this same category φ by means of a distinct construction Sy .
- c. Performance attrition in RL-dominant speakers causes them to express φ in the RL by means of a novel construction Ry , modelled on Sy , which does not reflect their underlying RL competence.
- d. Ry forms part of the PLD for children acquiring the RL, who thus acquire Ry as part of their underlying L1 competence.

In the case of Berber negation, the RL is Berber. The grammatical category φ is sentential negation. Construction Rx is the original Berber stage I negative construction with preverbal *ur* alone. The SL is Arabic. Construction Sy is the Arabic stage II negative construction with *mā...-š*. Construction Ry is the Berber bipartite

negative construction (taking on a number of different forms in the different varieties as shown in (22)-(25)). In this example the crucial stage (88)c presumably occurred roughly as follows. Berber-dominant speakers make extensive use of Arabic generally and thus also of the Arabic bipartite negative construction. This construction maps neatly onto the Berber construction, but with the addition of a postverbal negator which is (virtually) homophonous with an Arabic word meaning '(any)thing'.⁶¹ In their L1 Berber performance these speakers therefore recreate the Arabic bipartite construction by also adding a postverbal negator homophonous with a word meaning '(any)thing'. It is possible that in some instances this was done consciously and intentionally (the speakers having "noticed" the form of the Arabic construction), but this hardly seems a necessary condition for (88)c to occur. In any case, once this construction has made it into these speakers' L1 Berber performance, the possibility of children acquiring this construction as part of their L1 Berber is opened up (88)d.

The scenario in (88) also captures the facts of the KSKo relativization strategy borrowed from Kannada, which was presented in (78)- (81) in the previous chapter (cf. Heine and Kuteva 2005: 94-96). In this case, the RL is KSKo. The grammatical category φ is relativization. Construction Rx is the original KSKo/Indo-Aryan relativization strategy shown in (78), which is introduced by a non-polysemous

⁶¹ As we have seen (see footnote 13), although in most stage II Arabic varieties today the postverbal element is the enclitic *-š*, not obviously identifiable in synchrony with a word *šī* or *šay?* '(any)thing', there remain to this day certain varieties where the postverbal element has not (in all cases) undergone this phonetic reduction to *-š*, and thus still surfaces as *šī* or *šey*. Thus it seems a reasonable assumption that this lack of reduction was more widespread in the period of transfer of this construction to the various Berber varieties (certainly before 1738 in the case of Kabyle at least – see example (45)) and that the Arabic postverbal element would therefore have been homophonous with the word for '(any)thing' in this period.

relative pronoun and has no clause-final particle. The SL is Kannada. Construction *S_y* is the Kannada relativization strategy shown in (79)a, which is introduced by a relative pronoun identical in form to a separate interrogative pronoun, and which features an obligatory clause-final suffix identical in form to the main clause interrogative complementizer. Construction *R_y* is the KSKo relativization strategy shown in (79)b, which is also introduced by a relative pronoun identical in form to a separate interrogative pronoun, and which also features an obligatory clause-final suffix identical in form to the main clause interrogative complementizer. The presence of two new elements in the contact-induced KSKo relativization strategy makes it harder to be confident about just how the performance attrition in (88)c manifested itself in this case. Let us assume, however, for the sake of simplicity, that the KSKo use of *-ki* as a main clause interrogative complementizer is the result of an earlier instance of contact-induced grammaticalization on the model of Kannada. This being the case, one possibility for how (88)c manifested itself in the case of the KSKo relative clause is as follows. In acquiring Kannada as an L2, KSKo-dominant speakers register (not necessarily consciously) that a single phonetic string, *yāva*, signifies both an interrogative pronoun and a relative pronoun, and that a second string, *-ō*, is found both as a main clause interrogative complementizer and in relative clauses in final position. Regular use of the Kannada relativization strategy results in the recreation of these equivalences in the RL performance of KSKo-dominant speakers, that is, KSKo speakers extend the use of their own interrogative pronoun, *khanco*, and main clause interrogative complementizer, *-ki*, to relative clause contexts. Again, this then opens up the possibility of this novel construction being acquired as part of children's L1 KSKo competence, as described in (88)d.

While it is perhaps relevant to make a distinction between those instances of contact-induced grammaticalization which do result in the creation of a new grammatical category and those which do not, it would be preferable from a theoretical point of view to provide a single characterization of all types of contact-induced grammaticalization. I suggest the following:

(89) Contact-induced grammaticalization

- a. The general cognitive goal of minimizing processing costs associated with bilingualism results in the recreation of an SL construction Sx in the RL performance of RL-dominant speakers.
- b. This new RL construction Rx does not reflect the underlying competence of the RL-dominant speakers, but does form part of the PLD for children acquiring the RL, who thus acquire Rx as part of their underlying L1 competence.

It seems to me that (89) achieves an integration of Heine and Kuteva's insights concerning contact-induced grammaticalization into the innatist framework for contact-induced syntactic change proposed here on the basis of Van Coetsem's RL/SL agentivity distinction. As such, I suggest that (89) can form the basis of a more explanatory account of the many examples of contact-induced grammaticalization that Heine and Kuteva (2003, 2005) identify, as well as any others that have yet to be identified. The performance attrition approach to borrowing is schematized in Figure 4.2.

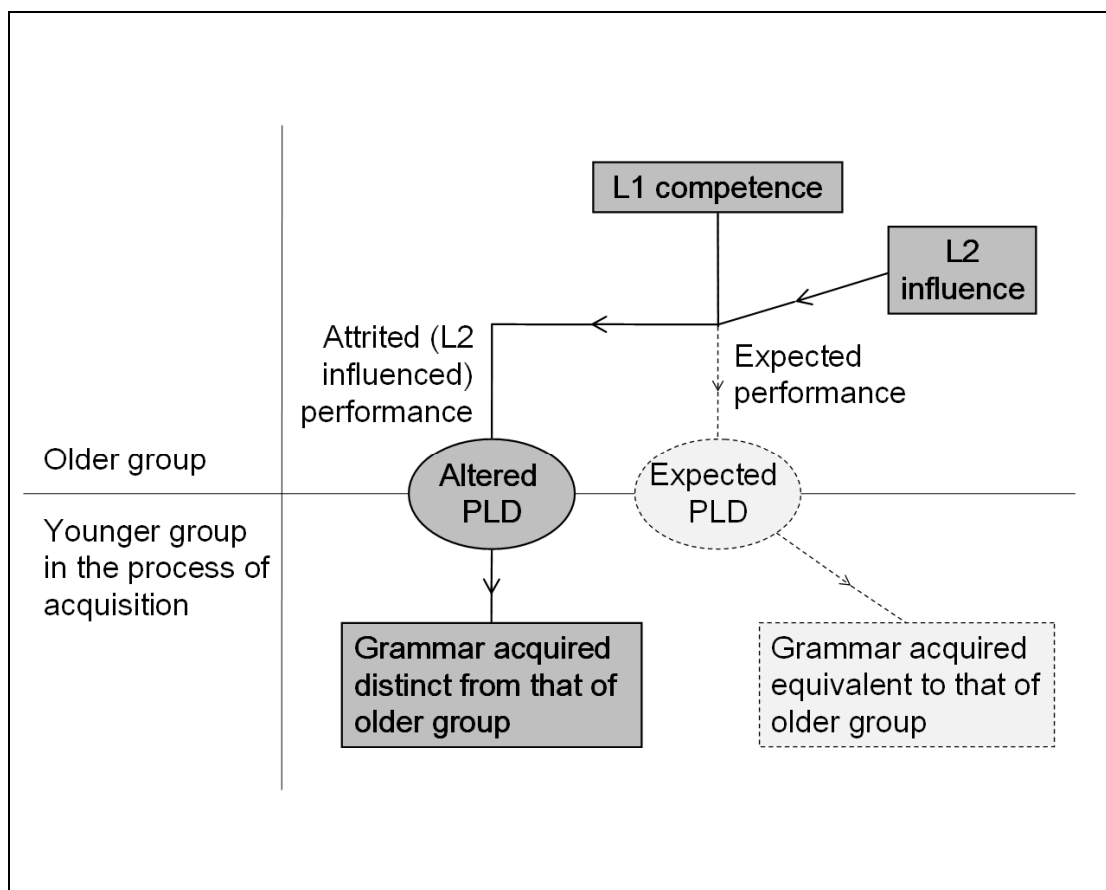


Figure 4.2

How borrowing leads to change

4.3 English restructuring

In chapter 3 it was argued that there is at least a *prima facie* case for the change from basic OV to VO order in English having involved contact with Old Norse. In common with probably every putative case of what Matras and Sakel (2007) call pattern transfer, it seems doubtful that one could provide evidence that constitutes actual proof either that contact was or was not involved here. There have already been several accounts of how this change might have occurred on the assumption that contact was not involved, as indicated in section 3.4.3. As long as there is at

least as much possibility that contact was involved, however, it is desirable to try to make such an account explicit.

The scenario that Weerman (1993) depicts for how this change actually took place is quite fine-grained and nicely illustrates some of the concepts we have been developing in this connection. In our terms he essentially claims (1993: 923) that the change was implemented by a combination of restructuring (L2 agentivity) and a form of indirect borrowing (L1 agentivity). This works as follows: L1 Norse speakers in the Danelaw imperfectly learn Old English and make the erroneous assumption that it has VO order, as is typical of L2 acquirers of languages that are OV with V2 (Clahsen and Muysken 1986). In some cases the VO tokens they produce make up part of the PLD for children acquiring English, who therefore acquire a competence with underlying VO order. This is thus a change due to restructuring (and not imposition – clearly, in this instance, because Old Norse appears to have been OV at this time; see Roberts 2007: 394 for a summary of the literature on Old Norse word order).

At the same time these Old Norse speakers will have been in contact with L1 speakers of Old English. This is likely to have resulted in some of the latter undergoing performance attrition due to frequent exposure to target-deviant strings from the L1 speakers of Old Norse. Note that this is somewhat different to simple borrowing as we have characterized it thus far: the L1 performance attrition of these Old English speakers occurs not as a result of having themselves acquired the SL as an L2, but rather through exposure to the interlanguage of L1 speakers of the SL. This seems plausible because, as noted in section 3.4.3, the VO order was already an option in Old English, albeit as a minor, marked pattern. This marked pattern

could thus have begun to be used more frequently and in a wider range of contexts as a result of persistent exposure to the VO L2 variety of English spoken by the L1 Old Norse speakers. Weerman (1993: 923) puts it as follows: “The relevant [L1 Old English] speakers do not change their internalized setting of the head parameter. From the perspective of their L1 grammar these overgeneralizations are ungrammatical. What they do is add a peripheral rule. A next generation, however, could set the head parameter differently.” In our terms it is performance rather than competence that is affected here. However, as Weerman implies, once these overgeneralizations have made it into the L1 performance of Old English speakers, this constitutes altered PLD for children acquiring the language and will therefore potentially lead to them acquiring a competence that has an altered basic word order relative to that of older speakers. At this point we can say that the change has taken place, due both to this indirect form of borrowing, and also to restructuring.

It is likely that this interplay between restructuring or imposition on the one hand and subsequent indirect borrowing on the other will be a feature of many changes which have their origins in SL agentivity. But again we have the same basic mechanism underlying these changes: an alteration in the PLD for children acquiring a language leads to the abduction of a grammatical competence that is distinct from that of older groups.

4.4 Summing up

The aim of chapters 3 and 4 has been to move towards a more explicit, psycholinguistically-based model of contact-induced syntactic change, in order to be

able to give an explicit account of the contact-induced spread of stage II negation from Coptic (and Modern South Arabian) to Arabic and from there to Berber.

One upshot of the discussion from this perspective is that it does not seem fruitful to try to provide a single unifying framework for contact-induced change across all linguistic domains – the mechanisms involved in lexical borrowing versus phonological imposition versus syntactic borrowing appear to be similar only at a trivial level. However, it is clear that contact-induced syntactic change can be integrated into a unified model for syntactic change in general, where this is based on Andersen's (1973) notion of change through abduction. Although the sociological factors that feed into contact-induced change are often so complex and varied as to lie beyond the scope of such a model, the change (innovation) itself need not and should not. If Lightfoot (1997: 269) is right that "there can be no change in grammars without change in trigger experiences", (or, as Longobardi 2001: 278 puts it, "syntactic change should not arise, unless it can be shown to be *caused*" (emphasis in original)), then both internal and externally-caused syntactic change must consist in children abducting their grammatical competence on the basis of a set of PLD which does not unambiguously reflect the competence of older groups of native speakers. The task of the historical linguist is then to account for how the trigger experience for these children came to be different from that of the older groups. One common way for this to happen is through contact with another language.

I have suggested here that we cannot understand how contact leads to change in trigger experiences without Van Coetsem's (1988, 2000) distinction between RL and SL agentivity. Speakers who have learnt an RL as a second language will typically produce utterances which are ungrammatical for native speakers of the RL. This

appears to be because the general learning strategies they use to acquire the RL differ markedly from the way in which children acquire their L1. Sometimes second language learners draw on some mental representation of their L1 and thus transfer elements of it onto the RL (imposition), and on other occasions their (mis)analysis of the RL leads to their producing utterances which are grammatical neither in the RL nor in their L1 (restructuring). In either case these hitherto ungrammatical utterances are likely to constitute part of the trigger experience for children acquiring the RL, potentially resulting in a change under SL agentivity.

By contrast, native speakers of an RL who have a high degree of exposure to an SL (including indirectly through the interlanguage of native speakers of the SL) will tend, in a variety of ways, to recreate elements of the syntax of that SL in their RL speech, resulting in utterances that are ungrammatical with respect to their own underlying competence. This appears to be because the (often involuntary) recreation of elements of their L2 in their L1 performance is a means of minimizing the processing costs associated with the extensive use of two languages. Again, if these utterances form part of the trigger experience for children acquiring the RL, this could result in a change at the level of competence, this time due to RL agentivity (borrowing). This, I suggest, is a useful way of understanding how contact leads to change in grammars.

5 Focus on Arabic

Having addressed the issue of the genesis of stage II negation in Arabic and its subsequent spread to Berber, the present chapter looks in more detail at the synchronic syntax and semantics of various types of negative sentences in contemporary spoken Arabic and offers a diachronic explanation for some of the patterns observed.

In section 5.1 we look in more detail at the contexts in which a stage I construction is either possible or obligatory in dialects where normally only a stage II (or stage III) construction would be possible. This raises the question of the (semantic) status of each negator in the bipartite construction and how they interact to supply a single semantic negation. I sketch an account of this in construction-based terms and indicate how the same account could be formulated from a Minimalist perspective. Section 5.2 focuses on n-words in a number of Arabic varieties, arguing that the varieties that exhibit true negative concord are fewer than has been claimed in the literature. Finally, section 5.3 offers an account of the progression to stage III in a subset of the Arabic varieties that have undergone Jespersen's Cycle, with a focus on Palestinian Arabic.

5.1 Residual stage I and the semantics of stage II negation in Arabic

We saw in section 2.3 that there are a number of contexts in which a stage I construction is either optional (in Libyan, Egyptian and Palestinian) or obligatory (Tunisian westwards) in dialects where a stage II construction is obligatory in

unmarked contexts. The stage I-favouring contexts already discussed are repeated here for convenience:

1. With 'never' expressions:

(90) **ʔomr-u ma-ža(*-š)**

age-his NEG-come.PRF.3MSG

'He never came.'

(Tunisian Arabic; Chaâbane 1996: 122)

2. With indefinite pronoun arguments:

(91) **Ma kienu jxewxu(*-x) l-ħadd.**

NEG AUX.PST.3PL provoke.IMP.F.3PL to-nobody

'They weren't provoking anyone.'

(Maltese; Sutcliffe 1936: 212)

3. Co-ordinate structures:

(92) **ma kla(*-š) ma šrəb(*-š)**

NEG eat.PRF.3MSG NEG drink.PRF.3MSG

'He neither ate nor drank.'

(Morrocan Arabic; Adila 1996: 108)

4. Oaths invoking God

(93) **wəllāh** ma-ngūl-ha-lu(*-š)

by God NEG-say.IMPF.1SG-it-to.him

'By God, I won't tell him it.'

(Moroccan Arabic; Caubet 1996: 86)

In addition, there are a number of further contexts for the optional or obligatory absence of the postverbal marker in some or all of the Maghrebi dialects (Moroccan, Algerian and Tunisian), as listed by Mettouchi (1996: 193). These include: a) when the complement of the verb is a fused relative construction (94), b) when the complement of the verb is an indefinite noun phrase (95), including minimizers (96), and c) with the adverb *gāf* 'at all', whether this precedes (97)a or follows the verb (97)b.

(94) **mā** šətt(*-š) **ləmmən** nəški

NEG find.PRF.1SG to-who complain.IMPF.1SG

'I haven't found anyone to complain to.'

(Algerian Arabic; Elhalimi 1996: 149)

(95) **ma** klīt(*-š) **xubz**

NEG eat.PRF.1SG bread

'I didn't eat any bread.'

(Tunisian Arabic; Chaâbane 1996: 123)

(96) ma kayswa(*-š) **bəšla**

NEG be.worth.IMPF.3MSG onion

‘It’s not worth an onion.’

(Casablanca Moroccan; Adila 1996: 112)

(97) (a) **gāf** ma ža(*-š) əl-yom

at.all NEG come.PRF.3MSG today

‘He didn’t come at all today.’

(b) ma tqəlləq(*-š) **gāf** ʔli-h

NEG grow.angry.PRF.3MSG at.all on-him

‘He didn’t get at all angry with him.’

(Casablanca Moroccan; Adila 1996: 106)

Accounting for these restrictions is, at least in principle, relatively straightforward from a diachronic perspective if one considers the likely extent of use of the Arabic stage II construction in its earliest days. I argued in the previous chapter that the development of the stage II construction could be said to have occurred in colloquial Egyptian Arabic at the point at which it became a feature of the grammars of L1 speakers of the language. Something being a feature of a grammar does not, of course, entail that there is any particular context in which that feature is obligatorily realized. Certainly the stage II construction is obligatory (and the stage I construction ungrammatical) in unmarked contexts in the majority of stage II Arabic varieties spoken today, but this is not in fact true of all of them. In urban Palestinian, for example, there appears to be full optionality between a stage I or stage II (or stage III – see section 5.3 for details) construction in unmarked contexts, and presumably

there has never been a time when the stage I construction has been ungrammatical in these contexts in this variety.⁶² It seems highly likely that for at least some time in the early period of the Arabic stage II construction there was no context in which it was obligatory and where the original stage I construction was ungrammatical. This being the case, two questions suggest themselves: what was the process by means of which the stage II construction became obligatory in unmarked contexts, and, at the point when it remained optional in all contexts, is it likely to have been less frequent in some than it was in others? It seems to me that the notion of emphasis is key to answering both of these questions.

Emphasis is a concept that is frequently appealed to in the literature (especially the literature on Jespersen's Cycle) but rarely defined. This is not the place to embark on an in-depth exploration of the topic. Let us instead follow Israel (1998, 2001) (see also Krifka 1995; Eckardt 2006: 151-155; and Schwenter 2006), who has provided a useful working definition of semantic and pragmatic emphasis as being when the proposition expressed entails what one might normally expect to be expressed, given the context. Thus, *I didn't sleep a wink* entails *I didn't sleep*. The emphatic value afforded by the perception of this entailment relation can be understood in terms of a general cognitive-pragmatic principle such as Sperber and Wilson's (1995) principle of relevance. On this view, given the 'presumption of optimal relevance' (Sperber and Wilson 1995: 270) a hearer is justified in deriving extra cognitive effects from a given utterance to the extent that the proposition

⁶² In this particular case, however, another possibility is that the stage I construction in unmarked contexts had in fact become ungrammatical or very rare in urban Palestinian, as it is in rural varieties, at some point in the past, and was then reintroduced as a prestige feature following contact with stage I varieties spoken in other urban centres in the Levant, such as Damascus.

expressed departs from a salient unmarked alternative. These cognitive effects could in principle be of a range of types, but in the case of emphasis they would amount to the assumption that the proposition in question holds more absolutely than one might otherwise have assumed.

In fact, this same principle extends to cases where the proposition actually expressed and what one might have normally expected are identical, and it is just the means of its expression that varies. This, by hypothesis, is the situation we had in Arabic after the stage II construction was developed but before it became obligatory in any context. A simple proposition such as ‘I didn’t sleep’ could be expressed by either (98)a or (98)b:

- (98) (a) mā nimt
 NEG sleep.PRF.1SG
 ‘I didn’t sleep.’
- (b) mā nimt šay?
 NEG sleep.PRF.1SG NEG
 ‘I didn’t sleep.’

Given that the whole of (98)b entails (is in fact identical in meaning to) the whole of (98)a, hearers could potentially interpret (98)b as more emphatic, as long as (98)a is felt to be the default or unmarked means of expressing the same proposition. This seems likely given the reasonable assumption that the stage II construction would have started out less frequent, and therefore more ‘extravagant’ (in the terms of

Haspelmath 1999; see also Kiparsky and Condoravdi 2006 and Detges 2003 on this point) than its original stage I counterpart.

Note that, according to the model of emphasis outlined above, emphasis can be, but is not necessarily, an inherent, semantic property of individual lexical items or constructions. This is an important distinction: between constructions whose use can trigger an emphatic interpretation (among other possible cognitive effects) in virtue of their being other than the norm, and lexical items which are semantically (rather than merely pragmatically) emphatic.

As we saw in section 2.1, apparently every language in the world has at its disposal elements which, when they co-occur with the ordinary expression of negation, necessarily serve to emphasize the negative polarity of the clause in which they appear. This is why it is necessary to stipulate that a genuine stage II construction has to be capable of being non-emphatic in at least some contexts. Without this stipulation one could justifiably claim that every language in the world had stage II negation. It is important to stress, therefore, that in the scenario developed here, emphasis is not part of the semantics of the postverbal marker (*šayʔ/šii/-š*); it is a possible pragmatic effect of the use of the stage II construction where the more frequent stage I alternative would have seemed sufficient.

We can illustrate this distinction with an English parallel. In (99)a we have a straightforward negative sentence. (99)b is identical except that negation is expressed not with the clitic *-n't* as would ordinarily be the case in an unmarked negative sentence in colloquial discourse, but with the full form *not*.

(99) (a) You can't do that.

(b) You cannot do that.

(99)b is perfectly grammatical in even the most colloquial speech, but its divergence from the unmarked alternative will, in most contexts, result in it being interpreted as more emphatically negative than (99)a. I take it as uncontroversial, however, that this does not mean that we should analyse the lexical item *not* as being inherently, semantically emphatic. To do so would require us to postulate two homophonous lexical items *not*, one emphatic, one not – a move that would clearly contravene Grice's (1978) Modified Occam's Razor in its lack of theoretical parsimony: "senses [i.e. distinct lexical items corresponding to a single phonological string – CL] should not be multiplied beyond necessity." This contrasts with a minimizing adjunct such as *one bit* in (100), whose semantics appears to be inherently emphatic, regardless of context.

(100) I don't like that one bit.

Hence *one bit* is not usually considered to be part of a bipartite negative construction; certainly not on the definition of bipartite negation adopted here.

Assuming, then, that optionality between *not* and *-n't* in English is a valid parallel for the optionality that would have existed between stage II and stage I negation in Arabic in the early days of that latter, sentences with a stage II construction would have been interpreted as emphatic in this period relative to a stage I alternative,

where the focus of emphasis would have been on the predicate in a sentence such as (98)b.

This being the case, we would not expect, generally speaking, that the stage II construction would be employed in contexts in which a) there is already some other means of emphasis, or b) where emphasis on the predicate is unwarranted for some reason. All of the contexts illustrated in (90)-(97), however, would appear to fit into one or other of these categories. ‘Never’ expressions (90), indefinite pronoun arguments (91), oaths (93), minimizer arguments (96) and *gāf* ‘at all’ (97), all straightforwardly fit into the first category: they form part of propositions that entail what one might normally expect to be expressed. Of course, one means of emphasis doesn’t preclude the possibility of a second being employed simultaneously, but it does render it unnecessary and therefore presumably less frequent.

Co-ordinate structures (92), fused relative construction complements (94) and ordinary, non-minimizer indefinite noun phrase arguments (95) fall into the second category. With co-ordinate negation, focus is on the contrast between the two negated elements (either predicates or arguments), or alternatively on the negative conjunction. This appears to prevent the possibility of an additional emphatic focus on the predicate(s). The fused relative construction cases, in which the complement of the verb is a noun phrase whose head is also the relative pronoun in a modifying relative clause, are probably best seen as a sub-type of the indefinite noun phrase argument cases, since the referents of these fused relative constructions seem always to be indefinite (this is true of all the examples given by Adila 1996, Chaâbane 1996, and Elhalimi 1996 to illustrate the phenomenon in question). The reason that emphasis on the predicate is unwarranted when an argument of that

predicate (typically an object) is indefinite seems to be that indefinite noun phrases tend to introduce referents that are discourse-new, and a listener is unlikely to have a normal expectation of what should hold with respect to a discourse-new entity, such that a proposition that entails this expectation can be expressed and accordingly be interpreted as emphatic.

Evidently there is a good deal more that could be said on all of these cases, but the essential point should be clear: on the reasonable assumption that there was a time when sentences featuring the stage II construction could be interpreted as emphatic with respect to counterparts with the stage I construction, the former is unlikely to have been much used in the contexts illustrated in (90)-(97). A child acquiring Arabic on the basis of PLD in which the stage II construction was largely absent in these contexts could then potentially learn this as the grammatical restriction we observe today.

We know, however, that the stage II construction eventually became obligatory in unmarked contexts, and the stage I construction ungrammatical in those same contexts, in most of the stage II varieties. This is probably best explained as resulting from inflationary use of the initially emphatic stage II construction, brought about by speakers' desire to mark their utterances out as particularly worthy of hearers' attention, and their resulting overuse of the construction in contexts where what the speaker has to communicate is not in fact especially worth attending to (cf. Detges and Waltereit 2002: 183; Detges 2003; Eckardt 2006: 166; Haspelmath 1999). Of course, the more the stage II construction is used in the contexts that allow it, the more the frequency of the stage I construction will tend towards zero in those same contexts, with the result that acquirers will be increasingly unlikely to acquire it as an

option in those contexts. Once there is no way of expressing negation in unmarked contexts relative to which the stage II construction can be understood as more extravagant, this construction will have become both obligatory in these contexts and entirely non-emphatic.

There does not appear to be any evidence for the alternative to this scenario, observable in French, West Flemish (Breitbarth and Haegeman 2009) and apparently Modern South Arabian (see example (63)), in which the rise of the stage II construction is accompanied by the simultaneous reanalysis of the original preverbal marker as a marker of non-assertive polarity rather than negation proper. The very fact that there are these contexts in all stage II varieties in which *mā* can be the only negative expression militates against an analysis of it as non-negative. An interesting red herring in this regard is the very common use of *mā* (or a reduced form *ma*) in Cairene and other stage II varieties as a non-negative emphatic particle, rather close in meaning to German *doch*, but difficult to translate into English:

(101) (a) **ma**-nti mrāt-i wi tiʔrabī-lu

EMPH-you.FSG wife-my and be.related.IMPF.2FSG-to.him

'You *are* my wife after all, and you're related to him.'

(Woidich 2006: 167)

(b) **ma**-tīgi rāyih fēn

EMPH-come.IMPF.IRR.2MSG go.PTCP.MSG where

'Come on! Where are you going?'

(Woidich 1968: 213)

Woidich (1968: 214) is probably right to suggest that this use of *mā* is the result of a reanalysis of a negative interrogative structure, used to express a strong affirmative declarative (compare English *Aren't you a clever boy!*). In any case, it cannot be that this reanalysis was triggered by the appearance of the postverbal negator, since this is a widespread feature of a range of Arabic dialects including stage I varieties such as Sudanese and Syrian (Reichmuth 1983: 111, Cowell 1964: 349), suggesting that this split of *mā* into two homophonous items, one negative the other emphatically affirmative, is a rather old innovation that predates Jespersen's Cycle.

This leads us to the question of the semantic status of each of the negative markers in stage II varieties of Arabic. If preverbal *mā* is negative, can it be that postverbal *-š* is as well? And, if so, how is it that the two items can co-occur in a single clause without cancelling each other out so that an affirmative sentence results? A denial of the inherent negativity of postverbal *-š* will be difficult to maintain. Apart from the fact that there are no obvious sources of evidence to support this position (as there are for French *ne*, for example), if we were to maintain on purely theoretical grounds that *-š* could not be negative in typical stage II Arabic varieties, we are left unable to account for the optional stage III construction found in Palestinian and other dialects, except again by making the unparsimonious and empirically unjustified stipulation that the *-š* in the stage III construction is a separate item to the *-š* in the stage II construction.

Similarly, it will not do to simply label the stage II construction a discontinuous but unitary morpheme (or a circumfix) whose two parts work together to provide a single semantic negation, because again we would encounter difficulties explaining how it is that in many cases the presence of just one of these two elements is sufficient to

make a clause negative. The empirical conundrum we are faced with is that either *mā* or *-š* alone are able to identify negation, whereas the two together do not result in a double negative, that is, affirmative, interpretation. As such, it seems preferable in cases such as these to see semantic negation as a property, not of individual lexical items, but of the whole proposition expressed by a sentence. That is, we should relinquish the assumption of rigid compositionality and acknowledge that the contribution of a given negative expression to the meaning of a sentence might vary according to the type of sentence, the expression's position within that sentence, and the potential co-occurrence of other negative expressions.

This makes sense from an acquisitional point of view, given that children must, in some sense, be expecting to find out how clauses are negated in the grammar they are acquiring; and they are clearly willing to tolerate the presence of a range of different syntactic means of conveying that the proposition expressed by an utterance contains a single logical negation. It is also incontrovertible that varying degrees of idiomaticity, that is, non-compositionality, are a pervasive feature of natural language, and must therefore be acquirable as part of grammars. In the specific case of negation there are sentences in English and many other languages which are interpreted as negative despite not containing any syntactic expression of negation, for example with the construction *fail to* in English, as in *I fail to see the relevance of that comment*.

This intuition also gives us a way of understanding the phenomenon of negative concord in general (of which stage II constructions can be seen as a special case), as we will see in more detail in the following section. Moreover, if formalized, this intuition is actually rather close in spirit to recent Minimalist approaches to the

problem of negative concord (e.g. Zeijlstra 2004, Penka 2007), although these approaches tend to see themselves as conserving the principle of compositionality with the tools of the formalism.

An outline of how one could formalize the present approach in Minimalist terms is as follows. The intuition that negation is primarily a property of whole propositions, to which individual negative expressions can contribute in varying ways, is equivalent to the postulation of an abstract (phonologically null) negative operator, usually symbolized as OP_{\neg} (e.g. Zeijlstra 2004). This operator carries an interpretable negation feature [iNeg], meaning that it is responsible for supplying a single logical negation in the logical form of the syntactic structure of which it is a part. The one or more overt expressions of negation that appear in a negative clause are assigned uninterpretable negative features [uNeg]. Thus they identify the presence of OP_{\neg} , whose [iNeg] feature is required to value their [uNeg] features, in order to allow the derivation to converge at Logical Form, but the overt negative expressions do not, themselves, each contribute a semantic negation to the proposition expressed. In these terms, then, the reanalysis of *šay?* as a negator involves the loss of its semantic content and the acquisition of a [uNeg] feature.

An approach along these lines captures the fact that either *mā* or *-š* or both together can express negation in certain Arabic varieties, without having to posit semantic ambiguity in either *mā* or *-š*. An approach that posited [iNeg] features on overt expressions of negation would have to posit one for both *mā* and *-š* in order to capture the fact that either can express negation alone, and then further posit a homophonous non-negative counterpart of either *mā* or *-š* in order to capture the fact that together they still express negation rather than double negation (= affirmation).

As discussed above, the latter approach appears to have little empirical support, in addition to being more stipulative and less elegant than the former.

Building on these ideas, the following section addresses the issue of n-words and negative concord in Arabic.

5.2 Developments in Arabic indefinites

This section is primarily concerned with the questions: does (any variety of) present-day spoken Arabic exhibit negative concord, and, if so, does this represent a change with respect to Classical Arabic? It will be seen that the answer to these questions is by no means straightforward. Much of the difficulty stems from the lack of widely agreed definitional criteria for what constitutes negative concord. We can point to languages, such as non-standard English, that are uncontroversially good examples of negative concord languages, but we lack a set of necessary and sufficient diagnostic conditions for more problematic languages, of which Arabic is certainly an example. A separate issue, which we touched on in the previous section, is, once we have identified a language as exhibiting negative concord, what is the best way to account for this phenomenon from a syntactic and semantic point of view? Clearly the stance we take on this issue will have an influence on how we analyse the relevant data.

Despite the small amount of previous literature on the topic of negative concord in Arabic (principally Moroccan Arabic) there is nevertheless a range of views and analyses. Benmamoun (1997, 2000, 2006) discusses the Moroccan Arabic determiner *ḥatta* 'even/any/no' and labels it a negative polarity item (NPI) without addressing the possibility of analysing it as an inherently negative item that

undergoes negative concord with the sentential negator. In contrast to this, Ouali (forthcoming) analyses *ḥætta* as semantically negative and, following Watanabe (2004), appeals to the syntactic mechanism of Agree to explain the lack of a double negation reading when it co-occurs with sentential negation. Finally, Hoyt (2005) follows Benmamoun's analysis of *ḥætta* as a non-negative NPI, and, for Palestinian Arabic, postulates two homophonous determiners *wala* '(not) even/(n)either/no/any', one an NPI, the other a negative quantifier.

My own analysis will differ from all of the above in some respect. Although I follow Benmamoun and Hoyt, and differ from Ouali, in seeing Moroccan *ḥætta* as not semantically negative, Benmamoun and Hoyt offer no explanation for *ḥætta*'s ability to appear in negative non-sentential utterances without an overt sentential negator. I will suggest an explanation for this below. I will also argue against Hoyt's ambiguity analysis of Palestinian *wala*. In general I will suggest that it is misleading to describe either Classical Arabic or the modern dialects as negative concord languages, although some individual items in some dialects could reasonably be seen as undergoing negative concord. We will see that the main exception to this generalization is Maltese, which does appear to be more straightforwardly describable as a negative concord language. Moroccan has a complex and heterogeneous negative indefinite system, some aspects of which undoubtedly display negative concord, while others, including *ḥætta*, arguably do not.

5.2.1 Issues in the terminology and analysis of indefinites

Given the aforementioned definitional problems, some further theoretical preliminaries are required before we can examine the data in detail. Most importantly, we need clear definitions of the relevant technical terms.

The definitions of 'NPI' and 'negative quantifier' are relatively straightforward and these terms are used fairly consistently in the literature. That said, establishing which of the two terms best describes a specific item in a given language is often as much a theoretical as an empirical matter, as we will see. In principle, however, it is agreed that negative quantifiers are inherently, semantically negative, while NPIs are not.

Classic, uncontroversial examples of negative quantifiers come from non-negative concord languages, for example Classical Latin *nemo* 'nobody', standard German *nichts* 'nothing', and, of course, standard English *nothing*, *nobody*. The fact that all of these items contain a morphological marker of negation (at least from an etymological point of view) makes their analysis as semantically negative especially uncontroversial, but containing such an overt morpheme is not usually thought of as a necessary feature of negative quantifiers.

NPIs on the other hand, while not themselves negative, are restricted to appearing in certain non-affirmative contexts such as negation, interrogatives and conditionals. Clear examples of these are provided by standard English *anyone*, *anything*. These are clearly not negative, they are grammatical in any of the above contexts, but they are ungrammatical in affirmative declarative sentences:

(102) *I saw anyone.

The literature which tries to provide a unifying account of NPI-licensing contexts is very large and cannot be treated in depth here. Krifka (1995) provides a useful summary of previous approaches, and a relevant concept that has become very influential in recent years is that of nonveridicality (Montague 1969; Zwarts 1995; Giannakidou 1998). As the name suggests, veridical contexts preserve the truth of a proposition embedded within them, while nonveridical contexts do not necessarily. Negative, interrogative and conditional contexts all seem to be nonveridical in this sense.

However, NPIs are very frequently licensed also in the context of comparatives, which are not obviously nonveridical (cf. Giannakidou 1998: 151-153). It seems that here we need another relevant semantic concept – that of downward entailment.⁶³ Downward-entailing contexts reverse the entailment relations that hold between pairs of propositions, and as such there is a large degree of overlap between nonveridical and downward-entailing contexts. For example, negation can be shown to be downward entailing as follows: *John reads magazines* entails *John reads*, whereas *John doesn't read* entails *John doesn't read magazines*. Comparatives can be shown to be downward entailing in the same way (cf. von Stechow 1984: 29): *an apple is tasty* entails that *an apple is tasty or an orange is tasty*, whereas *a banana is tastier than an apple or an orange* entails that *a banana is tastier than an apple*, provided that *or* is interpreted as inclusive, not exclusive disjunction in each case.

Some NPIs have further restrictions, however. The subset of NPIs that are only licensed in the context of negation are referred to as strong NPIs (NPIs which are not

⁶³ On the other hand, it does not seem to be possible to dispense with the concept of nonveridicality as a licensing context for NPIs since interrogative contexts are not downward entailing.

strong may be called weak; Zwarts 1998). We should also note here Hoeksema's (1994) term 'semi-NPI', which refers to items that may occur in veridical, upward-entailing contexts but which are more frequent in the context of negation. We will see that a number of items in the Arabic varieties examined here have this property.

A third term, originally due to Laka (1990), is 'n-word'. This tends to be used less consistently than 'NPI' and 'negative quantifier'. For clarity, I follow here Giannakidou's (2005: 328) definition (almost)⁶⁴ precisely:

(103) N-word:

An expression α is an n-word iff:

- (a) α can be used in structures containing sentential negation or another α -expression yielding a reading equivalent to one logical negation; and
- (b) α can be interpreted negatively in non-sentential utterances lacking a negator.

It should be clear that, according to this definition, the set of n-words does not form a natural class. Rather, 'n-word' is a label that is applied to items which are problematic in that they have one property that is consistent with their being NPIs and another that is consistent with their being negative quantifiers. Once an item is identified as an n-word according to the definition in (103), therefore, it is still in need

⁶⁴ I have rephrased the second clause of Giannakidou's (2005: 328) original definition "α can provide a negative fragment answer" in terms that are both more general (n-words can appear in non-sentential utterances which are not answers to wh-questions) and more theory-neutral (we needn't necessarily analyse non-sentential utterances as having undergone ellipsis).

of an analysis as to how it manages to exhibit these apparently contradictory properties.

This understanding of ‘n-word’ is in keeping with Laka’s (1990: 107-109) original use of the term, which was as a label for the Spanish indefinites *nadie* ‘n.one’ (= ‘anyone/no one’), *nada* ‘n.thing’, *ningún* ‘any/no’, *nunca* ‘(n)ever’ and various others. The label ‘n-word’ was simply chosen to highlight the fact that many of these items in Spanish (as well as parallel items in Italian, Portuguese and many other Romance varieties) begin with /n-/. Importantly, it is no part of the meaning of this term that an n-word necessarily contains a morphological expression of negation, as many people who are new to the term assume. Laka (1990: 108) makes this clear when she points out that *nadie* and *nada* originate in *(homines) nati* ‘born men’ and *(res) nata* ‘born thing’, respectively, and have thus never been morphologically negative. Instead, like Giannakidou, Laka sees n-words as a puzzle requiring an analysis, and this is the approach I will pursue here.

Subtly different interpretations of the term ‘n-word’ can be found in the literature, however. I will not explore these in depth here, but let us note, for example, statements such as “n-words differ from negative polarity items in at least three ways” (de Swart 2006: 203), or “[the] non-negative approach [...] makes it impossible to distinguish between negative polarity items and n-words” (de Swart 2008: 2). These statements presuppose that the set of NPIs and the set of n-words share no members. The approach adopted here, based on the definition in (103), does not presuppose this. In fact I will argue that some, but not all, of the n-words in the Arabic varieties to be examined here are indeed best analysed as NPIs, while others are best analysed as negative quantifiers. Many previous analyses of n-words in other

languages do not adopt this kind of heterogeneous approach, arguing instead for a unified account of all n-words according to one of three analyses outlined below. However, I see no *a priori* reason to expect that all n-words in all languages should be of a single semantic kind (cf. Giannakidou 2006, who argues that Greek n-words are best analysed as NPIs, while Portuguese n-words are best analysed as negative quantifiers). When we consider the behaviour of Arabic n-words in contexts other than those specified in (103) (i.e., outside of sentences containing sentential negation or non-sentential utterances) we will see that the data point to a negative quantifier analysis in some cases and an NPI analysis in others. Thus, while our analysis of n-words must be theory-driven to a certain extent, there are also compelling empirical considerations which favour the range of analyses presented here.

The three main types of analysis of n-words, then, are as follows.

First, we can posit semantic ambiguity: a given n-word is in fact a pair of homophonous items, one an NPI, the other a negative quantifier. Examples of this type of analysis can be found in the work of Zanuttini (1989) and Herburger (2001). Such an analysis reduces negative concord to an epiphenomenon of this ambiguity. Although this appears to be a welcome result, positing homonymy must be seen as a last resort, given the associated risk of multiplying entities beyond necessity and thus contravening (Modified) Occam's Razor.

Second, we can analyse a given n-word as a negative quantifier (inherently negative).⁶⁵ This entails that negative concord is a real phenomenon, which itself then requires further explanation. Typically this is done by appeal to some technical

⁶⁵ As we will see, some n-words, such as Palestinian or Egyptian Arabic *lissa* '(not) yet/still', have no quantificational force, meaning they cannot properly be referred to as (potential) negative quantifiers, though they could in principle still be analysed as inherently negative.

device within a particular formal syntactic or semantic framework that allows for the ‘absorption’ of multiple semantic negations into one (e.g., Haegeman and Zanuttini 1991; de Swart and Sag 2002; Watanabe 2004). An alternative version of this approach is to relinquish the assumption of rigid compositionality, as in the discussion of the semantics of stage II negation in the previous section, and to say that the contribution of a negative quantifier to the meaning of a sentence might vary according to the type of sentence, and the negative quantifier’s position within that sentence (e.g., Jespersen 1924: 331-334, and, implicitly, van der Auwera and Neuckermans 2004). A cynic might view the absorption approach as the non-compositional approach in disguise.

Third, we can analyse a given n-word as an NPI (e.g., Laka 1990; Giannakidou 2000). This analysis also denies the existence of negative concord as a phenomenon in need of explanation. It does, however, raise the problem of how to account for the negative interpretation of such items in non-sentential utterances without an overt negator – clause (b) of the definition in (103). A reason to think that this might nevertheless be a valid approach in at least some cases comes from the English word *either*.

Either, in the sense of *I don’t like you either* (and not *either...or*), seems to be a fairly straightforward case of a strong NPI (modulo its scopal properties, which we will not deal with here). As shown in (104), it is grammatical in the scope of negation (as well as adverbs such as *hardly*, *rarely* etc.), but not in non-negative nonveridical contexts, or in veridical contexts, and it appears to be identical in meaning, though not in distribution, to *too*.

- (104) (a) I don't like fish, and I don't like chips either/*too.
 (b) Do you like fish, and do you like chips too/*either?
 (c) If you like fish, and you like chips too/*either, ...
 (d) I like fish, and I like chips too/*either.

On the basis of these data there is clearly no temptation to say that *either* is an n-word, or that it undergoes negative concord. However, there is a variety of English, which appears to be rare in Britain but fairly widespread in America, in which a) there is no generalized negative concord, just as in standard English, b) the grammaticality judgments of (104) hold, but c) *either* clearly is an n-word according to the definition in (103), since it can be, and regularly is, interpreted as negative in non-sentential utterances without an overt negator as in (105).

(105) A: I don't like chips.

B: Me either. (= Me neither = I don't like chips either.)

Does this mean that, unlike in standard English, *either* is inherently negative in the variety in which (105) is grammatical, and that just this item, but no other in this variety, undergoes negative concord? This seems a rather inelegant way to approach the problem. In this instance it seems preferable to develop an analysis of the non-sentential utterance such that here, as elsewhere, the negative context licenses the use of *either*, which is then analysed as an NPI also in this variety. This is the approach that Giannakidou (2000, 2006) takes with n-words in Greek and a number

of other languages with the appearance of negative concord, arguing that the semantic negation which licenses the n-word in examples such as (105) has undergone ellipsis. This is also the type of analysis that I will defend here for the majority of n-words in Classical/Modern Standard Arabic (MSA), Palestinian and Egyptian.

Note, however, that we need not necessarily analyse non-sentential utterances such as the one in (105) as fragments of full sentences, the remainder of which has undergone ellipsis. Rather, following Stainton (2005), Progovac (2006) and others, I propose to take them at face value and analyse precisely those utterances, rather than some partly unpronounced fully sentential form, as communicating the proposition recovered by the hearer. Giving a detailed sketch of such an approach is beyond the scope of this section, but it seems reasonable to suppose that languages might vary with respect to the set of contexts in which certain non-negative items can appear in non-sentential utterances lacking a negator but interpreted as negative. For example: no items in any context (standard English), just *either* in the context of the “*me + conjunction*” construction (a variety of American English), indefinites found predominately in negative contexts (Arabic, I will claim), and so on.

With this background in place, we can now turn to the Arabic data and provide an analysis. The varieties to be examined are: Classical Arabic/MSA (section 5.2.2), Palestinian/Egyptian (section 5.2.3), Moroccan (section 5.2.4) and Maltese (section 5.2.5).

(107) (a) lam ara aḥadan
 NEG see.JUSS.1SG anyone.ACC
 'I didn't see anyone.'

(b) hal raʔayta aḥadan
 Q see.PRF.2MSG anyone.ACC
 'Did you see anyone?'

(108) lam yaʕud qaṭṭu
 NEG return.JUSS.3MSG ever
 'He never returned.'

We may also note here the existence of an adverb *baʕadu* 'yet, still', which appears predominately in the scope of negation (109)a, but not exclusively (109)b, thus making it non-polarity-sensitive, albeit with significant strong NPI tendencies (a 'semi-NPI in Hoeksema's (1994) terms).

(109) (a) lam yaʕil baʕadu
 NEG arrive.JUSS.3MSG yet
 'He hasn't arrived yet.'

(b) huwa baʕadu ʕayīr
 he yet small

'He's still young.'

(Wehr 1979: s.v. *baʕadu*)

There is also a marker of exclusive disjunction *ʔam* (110) that is restricted to yes-no questions (hence a restricted kind of NPI) and a negative conjunction *walā* (111).⁶⁷

(110) hal tuṛīdu qahwatan **ʔam** šāyan
 Q want.IMPF.2MSG coffee.ACC or tea.ACC
 ‘Do you want coffee or tea?’

(111) (a) lan tuḡniya ʕan-hum amwālu-hum **wa-lā**
 NEG avail.IRR.3FSG from-them riches.NOM-their and-NEG
 awlādu-hum šayʔan
 children.NOM-their thing.ACC
 ‘Neither their riches nor their children will avail them anything.’
 (Qur’an 3: 10)

(b) ʔinna-hā baqaratun lā fāriḡun **wa-lā** bikrun
 COMP-it cow.NOM NEG old.NOM and-NEG virgin.NOM
 ‘It is a cow that is neither old nor immature.’
 (Qur’an 2: 68)

⁶⁷ We may also note here several other conjunctions closely linked with negation: *lā* ‘not’, *bal* ‘rather, in fact’ and *ʔillā* ‘except’. Conjunction *lā*, not to be confused with *walā* ‘nor’ is a constituent negator usually negating noun phrases or adjective phrases which contrast with a non-negated alternative, e.g. *b-il-ʔinklīziyyati lā b-il-ʕarabiyyati* ‘in English, not Arabic’. *Bal* has the inverse function of *lā* in that it instructs the hearer to infer a contrast between a negated constituent which precedes it and a non-negated constituent which follows, e.g. *mā samīʕtu ḡissan bal jarsan* ‘I didn’t hear a voice, but rather a bell’. Finally *ʔillā* is frequently found in the context of negation, where the meaning ‘not... except’ is usually best translated as ‘only’, e.g. *mā jāʔa ʔillā axī* ‘only my brother came’ (lit. ‘didn’t come [anyone] except my brother’).

This negative conjunction *walā* is important for two reasons. First, it undergoes some important developments into present-day spoken Arabic, as we will see shortly. Second, it could perhaps be viewed as an n-word. Although it appears not to occur in non-sentential utterances, it is transparently morphologically negative, containing as it does the negator *lā*. Despite this, it appears to satisfy the first clause of the definition of n-words in (103), particularly if we consider examples such as (111)a. Here *walā* appears to take scope over just *awlāduhum* ‘their children’, making this and *amwāluhum* ‘their riches’ a conjoined subject, which in turn is clearly in the scope of the sentential negator *lan*. Before we consider what the best analysis of this is, let us first note that something very similar can be observed in standard English. For example, from the writings of Charles Darwin:

(112) (a) ...for neither earth nor ocean could hold the product.

(b) ...neither transplantation or mutilation prevented the seed from being perfected.

(Stauffer 1975: 83, 180)

Both (112)a and (112)b are interpreted such that the predicate fails to hold of two referents. In (112)b *neither* alone clearly signals semantic negation, suggesting that it is also interpreted as negative in (112)a. How is it, then, that (112)a, which contains a second negative element *nor*, fails to trigger a double negation reading? Does *nor* undergo negative concord here? Again, given that negative concord is absent from any other domain of standard English grammar, it seems preferable to argue that it does not obtain here either. Instead one would probably argue concerning (112)a

that *neither earth* expresses one (negated) proposition, either as part of a fuller, ellipsed structure or simply in and of itself, while *nor ocean could hold the product* expresses a second (negated) proposition, with *neither* and *nor* functioning both as conjunctions and negators. (112)b, by contrast, is probably best analysed as involving a single negation taking scope over two conjoined propositions. (For more on the history of English *nor* see Ingham (forthcoming), who accounts for the lack of a double negation interpretation in examples such as (112)a by the presence of a [uNeg], rather than [iNeg] feature on *nor* in such contexts).

To the extent that this or some other argument for why English *nor* does not undergo negative concord is accepted, precisely the same case can be made for the directly parallel Arabic *walā*. Thus it seems reasonable to say that *walā* is negative and yet Classical Arabic/MSA sentences such as in (111) do not instantiate negative concord. This point will be important when we consider the developments of *walā* in present-day spoken Arabic dialects.

Classical/MSA can form genuine negative quantifier phrases by means of the constituent negator *lā* and an appropriate (pro)noun. These phrases are not n-words, since, like English *nothing/nobody*, they cannot be used in structures containing sentential negation to yield a reading equivalent to one logical negation (clause (a) of (103)). If they are used in such structures the result is ungrammaticality, or perhaps marginally double negation (113)a,b. These negative quantifier phrases are also only possible in preverbal subject position (113)c.

(113) (a) lā aḥada (*mā) jāʔa

NEG one.ACC NEG come.PRF.3MSG

‘No one came.’

(b) lā šayʔa (*mā) ḥadatha

NEG thing.ACC NEG happen.PRF.3MSG

‘Nothing happened.’

(c) (mā) ḥadatha šayʔun/ *lā šayʔa

NEG happen.PRF.3MSG thing.NOM NEG thing.ACC

‘Something (/nothing) happened.’

The only other candidate for a (weak) NPI is *ʔayy* ‘any’. Like English *any* it can also be used as a free-choice item in veridical contexts, but unlike English *any*, in nonveridical contexts it is always optional and emphatic, and does not have an effect on the aspectual interpretation of the predicate.⁶⁸

⁶⁸ That is, unlike English *any*, it cannot be seen as the NPI equivalent of the determiner *some* [sm], not least because, unlike English NPI *any*, it is compatible with singular count nouns:

(i) lam ataqabbal **ʔayya** hadīyatin

NEG receive.JUSS.1SG any.ACC gift.OBL

‘I didn’t receive a single gift / any gifts at all.’

- (114) (a) hal iṣṭarayta kutuban
 Q buy.PRF.2MSG books.ACC
 ‘Have you bought (any) books?’ [Note aspectual effect of (absence of) English *any*]
- (b) hal iṣṭarayta **ʔayya** kutubin
 Q buy.PRF.2MSG any.ACC books.OBL
 ‘Have you bought *any* books (at all)?’

As such it is unclear whether one should see *ʔayy* as a single, non-polarity-sensitive item, or as two homophonous items, one free choice, the other an NPI, as is often postulated for English *any*.

The rather meagre indefinite system of Classical Arabic/MSA is summarized in Table 5.1. (Blanks represent absent forms. Parentheses represent forms whose placement in a cell is doubtful, either because they are better classified differently, or because they are non-lexicalized phrases. Question marks represent possible analyses of n-words).

Table 5.1 Indefinites in Classical Arabic / MSA

	negative	n-word	strong	weak
	quantifier		NPI	NPI
Determiner	lā	—	—	(ʔayy)
Thing	(lā šayʔ)	—	—	(ʔayy šayʔ)
Person	(lā aḥad)	—	—	aḥad
Extent adverb	?	← abadan →?		—
Future time	?	← abadan →?		—
Past time	—	—	qaṭṭu	(qaṭṭu)
still/yet	—	—	(baʔadu)	—
Conjunction	walā	← (walā)	—	ʔam
Place	—	—	—	(fi ʔayy makān)

Having identified *abadan* as an n-word, further analysis is required to decide whether it is a negative quantifier that undergoes negative concord or a non-negative item capable of appearing in negative non-sentential utterances without a negator. There are two reasons for thinking that the latter option is the better analysis. First of all, we have seen that in Classical/MSA those phrases which are clearly, uncontroversially interpreted as negative quantifiers do not undergo negative concord. Secondly, it is not clear that *abadan* is a lexicalized indefinite. It is transparently formed from the word *abad* ‘eternity’ and the adverbializing suffix *-an*. As such, it is not surprising that it is also found in affirmative sentences meaning ‘forever, for all eternity’:

- (115) wa-jannātin la-hum fi-hā naʿīmūn muqīmūn
 and-gardens.ACC for-them in-them comfort.NOM permanent.NOM
 xālidīna fi-hā **abad-an**
 endure.PTCP.ACC.PL in-them eternity-ADV
 ‘...and gardens for them in which they will have permanent comfort, enduring
 in them forever.’ (Qur’an 9: 21-22)

Note that in both negative and affirmative sentences *abadan* takes scope over the whole proposition: ‘it is forever the case that (it is not the case that) *P*’. Clearly *abadan* in (115) is not negative, and yet its semantic contribution is apparently identical to that of *abadan* in (106). Positing two distinct homophonous items with identical meanings apart from their polarity, just to account for the use in (106)b, would appear again to fall foul of (Modified) Occam’s Razor. I therefore conclude that *abadan* is best analysed as a single polarity-neutral item (albeit a semi-NPI with a strong preference for negative contexts, like *baʿadu* ‘still, yet’ above), and that there is no strong evidence that would lead us to call Classical/MSA a negative concord language, despite the presence of one or more n-words.

5.2.3 N-words in Palestinian/Egyptian

Palestinian and Egyptian are dealt with together in this section because their respective indefinite systems are virtually identical.

In his discussion of the Palestinian indefinite system, Hoyt (2005) takes it as self-evident that this variety (and therefore, by extension, also Egyptian) exhibits negative concord. However, this appears to follow, from Hoyt’s perspective, from the fact that

Palestinian has n-words, the details of which are presented below. As should be clear from the above, this does not automatically follow from the definition in (103), and the approach taken here is that n-words are problematic items which may or may not be analysed as inherently negative.

Let us start, however, with (what I take to be) the counterfactual assumption that Palestinian and Egyptian are negative concord languages, as Hoyt (2005) suggests. We are then confronted with the fact that there are items which appear to be 'too negative' to undergo negative concord as expected. These items are *maḥaddiṣ* 'no one' (Palestinian and Egyptian) and *mafīṣ* 'nothing' (Egyptian). Both of these words are lexicalized negative pronouns which are clearly morphologically negative. We will see that this marks them out as distinct from almost all other n-words in the present-day spoken Arabic dialects. In the case of *maḥaddiṣ*, this is self-evidently derived from the weak NPI indefinite *aḥad* 'anyone', which we saw in the discussion of Classical/MSA above, plus bipartite *mā...-ṣ* negation. *Mafīṣ* is derived from the existential expression *fī* plus the same negative construction ('there is not' > 'nothing'). Both these items are rather restricted in their distribution: *maḥaddiṣ* can only occupy the subject position and must stand before the verb (116)-(117), while *mafīṣ* appears not to be able to occupy verbal argument positions at all (118)-(119), meaning that the logical negation it provides can never take wide scope over the whole proposition. Nevertheless, both *maḥaddiṣ* and *mafīṣ* are clearly inherently negative and when they co-occur with sentential negation double negation results (i.e. they are not n-words by (103)a):

- (116) **maḥaddiṣ** aja
 no.one come.PRF.3MSG
 ‘No one came.’ (Palestinian)
- (117) maḥaddiṣ min al-bašar **ma-lu-š** maḥāsin
 no.one from the-mankind NEG-have.3MSG-NEG good.qualities
 ‘No one in existence doesn’t have some good qualities.’
 (Egyptian; Woidich 1968: 73)
- (118) wa-māḏa kull hāḏihi š-šawšara ʕala **mafiš**
 and-what all this the-noise about nothing
 ‘What’s all this fuss about nothing?’
 (Educated Egyptian Arabic; <http://arabicblog.swissinfo.ch/?p=63>)
- (119) in-naḍḍarāt dōl **miš** aḥsan min **mafiš**
 the-glasses this.PL NEG better than nothing
 ‘These glasses are not better than nothing.’ (Egyptian)

So the only words in these varieties about which there can be no doubt that they are negative quantifiers do not participate in negative concord. This must raise doubts concerning any analysis of the n-words in these varieties as negative quantifiers rather than NPIs.

Where *maḥaddiṣ* and *mafiš* cannot be used, the meanings ‘no one’ and ‘nothing’ are expressed by sentential negation marked on the verb combined with a non-negative indefinite (cf. the Classical/MSA example in (113)c), such as the weak NPI *ḥadd* (Egyptian) or *ḥada* (Palestinian) ‘anyone’, or non-polarity-sensitive *ḥāga* (Egyptian) or *išī* (Palestinian) ‘thing’:

- (120) (a) **ma** šaf-nī-š **ḥadd/ḥada**
 NEG see.PRF.3MSG-me-NEG anyone
 ‘No one saw me.’
- (b) **ma** šuft-iš **ḥāga/iši** (ḥadd/ḥada)
 NEG see.PRF.1SG-NEG thing anyone
 ‘I didn’t see anything (anyone).’
- (Egyptian (/Palestinian); Woidich 2006: 337)

Palestinian and Egyptian have n-words only in the categories of determiner, extent adverb and ‘still/yet’ words: *wala* ‘(not) even a’ (both Egyptian and Palestinian), *abadan* ‘in no/any way’ (both), *b-il-marra* ‘in no/any way’ (lit.: ‘in-the-time’; Palestinian only), *lissa* ‘still/yet’ (both) and *baṣad* ‘still/yet’ (Palestinian only).

The last category is perhaps the most straightforward to analyse. The Classical/MSA adverb *baṣadu* is maintained in Palestinian *baṣad*, although it now predominately appears sentence-initially and it has been partially reanalysed so that it can function as an impersonal verb which takes the logical subject of the sentence it appears in as a pronominal object. *Lissa* (< *l-is-sāfa* ‘to the (current) time’) is very similar to *baṣad* in its meaning, function and distribution. Importantly, neither word shows any polarity sensitivity, and they are both clearly non-negative in the absence of sentential negation. Nevertheless they both satisfy clause (b) of (103):

(121) (a) hiyya **lissa** (/baʕad-ha) txīna

she still (still-her) fat.F

‘She’s still fat.’

(b) ʕayyib bass ana **lissa** (/baʕad-ni) ma-šuft-iš il-ṣarūsa

ok but I still (still-me) NEG-see.PRF.1SG-NEG the-bride

‘Ok, but I haven’t seen the bride yet.’

(Egyptian (/Palestinian); Woidich 2006: 167, 349)

(c) A: huwwa mayyit

he dead

‘Is he dead?’

B: **lissa** (/baʕad-u)

still (/still-him)

‘Not yet.’

(Egyptian (/Palestinian))

Given these facts, there seems to be no need to analyse *lissa/baʕad* as (potential) negative quantifiers. Rather, they are polarity-neutral items whose frequent association with negation allows non-sentential utterances which contain them to be interpreted as negative, provided the context is conducive to such an interpretation.

The properties of Palestinian/Egyptian *abadan* (and Palestinian *b-il-marra*) are similar to those of Classical/MSA *abadan* and the same arguments against seeing it as a negative quantifier, even when it appears in non-sentential utterances interpreted as negative as in (122)a, still apply. A development with respect to Classical Arabic is that *abadan* is now used principally as an extent adverb (122)b, rather than a temporal adverb (but see (122)c); the meaning of ‘ever’ has in general

been taken over by a construction with *ʕumr* < ‘age’, which is not an n-word). *Abadan* can also occur in interrogative clauses such as (122)c, where it is clearly non-negative, but it still seems to occur predominantly in the scope of negation (Woidich 2006: 349).

(122) (a) inni saḥbit-na tiskut, **abadan!**

COMP friend-our be.silent.IMPF.IRR.3FSG ever

‘That our friend would keep quiet? Never!’

(b) di masʔala miš sahla **abadan** (/b-il-marra)

this.F issue NEG easy.F ever (/ever)

‘This is an issue which is not at all easy.’

(c) huww-anta maʕā-na **abadan**

Q-you.MSG with-us ever

‘Do you ever agree with us?!’

(Egyptian/Palestinian; Woidich 2006: 162, 349)

Somewhat more problematic is the n-word determiner *wala* ‘(not) even a’.⁶⁹ There cannot really be any doubt that *wala*, unlike the n-words we have seen so far, should be analysed as a negative quantifier. A number of pieces of evidence taken together strongly suggest this analysis. First, *wala* is always interpreted as negative in non-sentential utterances without a negator (123). It also can occur in preverbal subject

⁶⁹ For an in-depth analysis of ‘even’-type words in Greek and other languages, see Giannakidou (2007).

position, in the absence of a sentential negator, again always being interpreted as negative (124).

- (123) **wala** kilma
 not.even word
 ‘Don’t say a word!’ (Egyptian/Palestinian; Woidich 2006: 342)

- (124) **wala** taksi wiʔif
 not.even taxi stop.PRF.3MSG
 ‘Not a single taxi stopped.’ (Egyptian/Palestinian; Woidich 2006: 342)

Furthermore, *wala* can occur in positions where it is not a verbal argument and therefore cannot contribute wide-scope negation to the proposition expressed by the utterance in which it occurs (125). That is, *wala* can appear in affirmative sentences, where it still contributes a (narrow-scope) semantic negation:

- (125) (a) šwayya aḥsan min **wala** ḥāga / iši
 little better than not.even thing / thing
 ‘A little is better than nothing at all.’ (Egyptian/Palestinian)
- (b) huwwa zaʕlān ʕala **wala** ḥāga / iši
 he angry about not.even thing / thing
 ‘He’s angry about nothing at all.’ (Egyptian/Palestinian)

This is significant for three reasons. First, it makes it impossible to claim that the semantic negation associated with *wala* is provided by a marker of sentential

negation, rather than by *wala* itself, because naturally there is no marker of sentential negation in an affirmative sentence. Second, unlike the Greek n-words in negative fragment answers that Giannakidou (2000, 2006) analyses, the examples in (125) are not amenable to an analysis which postulates ellipsis of a negator as part of some larger constituent, since it is not at all obvious what material within the relevant PPs could be claimed to have undergone ellipsis, and why. Third, we can see in (125)a that *wala* is still interpreted as negative in the (non-negative) downward-entailing context of a comparative. This makes it more clearly negative than the classic n-words of the Romance languages, which are interpreted as non-negative in comparatives, as in Spanish *ninguno* ‘n.one’, for example:

- (126) María canta major que ninguno de vosotros
 María sing.PRES.3SG better than n.one of you.PL
 ‘María sings better than any of you.’
 (Spanish; Laka 1990: 113)

As such, the conclusion that *wala* is a negative quantifier seems inescapable, given the evidence reviewed so far.

The difficulty arises in that, when *wala* appears as part of a postverbal subject or (indirect) object, it must co-occur with the sentential negator and the sentence is then interpreted as containing just one logical negation (which is why *wala* must be classed as an n-word unlike, e.g., *maḥaddiṣ* ‘no one’).

(127) **miš** sāmīf **wala** kilma

NEG hear.PTCP.MSG not.even word

‘I can’t hear a single word.’ (Egyptian/Palestinian; Woidich 2006: 342)

Having analysed *wala* as a negative quantifier, must we analyse (127) as an instance of negative concord? Hoyt (2005) considers it self-evident that that is what we have here and argues that the form /wala/ is ambiguous between a negative existential quantifier and a plain existential quantifier, the latter conveniently occurring just in those cases where the former would result in an interpretation of double negation. But given the identity of the interpretations of *wala* in (127) and (123), aside, apparently, from their polarity, it would seem preferable to give a unified analysis of (at least this sense of) *wala* if possible. In fact, consideration of the diachrony of *wala* supports such an analysis. It clearly derives from Classical Arabic *walā*, and in fact retains the negative conjunctive function of its Classical ancestor:

(128) ma-ṣanda-hum-šī zō? **wala** tarbiyya

NEG-have-they-NEG taste and.NEG upbringing

‘They have neither taste nor manners.’ (Egyptian; Woidich 2006: 344)

The analysis of utterances containing Classical Arabic *walā* as expressing two conjoined, negated propositions (in the same way as sentences containing English *nor*) is equally applicable here, such that, in (128) at least, we can analyse *wala* as inherently, semantically negative without having to speak of negative concord. It then

seems reasonable to extend this analysis to apparent cases of non-negative determiner *wala* as in (127), which we could analyse as expressing:

(127)' I can't hear ~~anything~~, not even a word.

(127)'' $\neg\exists x (\text{hear} (\text{Me}, x)) \wedge \neg\exists y (\text{word} (y) \wedge \text{hear} (\text{Me}, y))$

On this analysis of (127), *wala* still conjoins two negated propositions, where the material in the second that is identical to that in the first is unexpressed (or ellipsed), but in this case there is also an unexpressed indefinite pronoun argument in the first conjunct.

Alternatively one could simply say that *wala* has been fully reanalysed as a negative determiner 'no(t even an) *x*', which is a separate item to the negative conjunction *wala*, and which, unlike the other negative quantifiers in Palestinian/Egyptian, undergoes negative concord. In this case, rather than positing two items that are identical in every respect apart from their polarity (as in Hoyt 2005, Herburger 2001), it would seem preferable to give an account of negative concord structures containing *wala* in terms of the relaxation of strict compositionality, along the lines suggested in the previous section. We could formalize this by saying that *wala*, in common with all overt negative expressions including the sentential negator(s), bears a [uNeg] feature that is valued by the [iNeg] feature on $\text{OP}\neg$. That is, the presence of *wala* identifies a clause as negative (which would not be the case

if it were an NPI), as does the sentential negator, but semantic negation is a property of the proposition as a whole, rather than of these individual negative expressions.⁷⁰

Table 5.2 summarizes the indefinite system of Egyptian (E) and Palestinian (P) Arabic. Arrows indicate the analyses of n-words given here. Parentheses indicate non-lexicalized items or items which are not polarity-sensitive but which are nevertheless frequently found in negative or nonveridical contexts.

Table 5.2 Indefinites in Egyptian and Palestinian Arabic

	negative	n-word	strong	weak
	quantifier		NPI	NPI
Determiner	wala ←	wala	—	(ʔayy) ^a
Thing	mafīš - E	—	—	(ʔayy ḥāga - E) (ʔayy iši - P)
Person	maḥaddiṣ	—	—	(ʔayy) ḥadd - E (ʔayy) ḥadā - P
Extent adverb	—	abadan →		abadan
		b-il-marra - P →		b-il-marra - P
Time	—	—	—	ʔumr-
still/yet	—	baʔad - P →	(baʔad - P) ^b	—
		lissa →	(lissa) ^b	
Conjunction	wala ←	wala	—	—
Place	—	—	—	(fi ʔayy makān)

^a As in Classical/MSA, *ʔayy* is also found in veridical contexts where it functions as a free choice item.

^b As noted above, *baʔad* and *lissa* are polarity-neutral but frequently occur in the context of negation ('semi-NPIs').

⁷⁰ In cases such as (125) where *wala* does not occupy a verbal argument position and the sentence as a whole is interpreted as affirmative, $OP\bar{\neg}$ would have to occupy a position within the PP, where it would be out-scoped by the predicate.

5.2.4 N-words in Moroccan

The situation in Moroccan Arabic is rather different.⁷¹ Here there are two n-words that merit particular attention: *wālu* ‘n.thing’ (whose etymology is not certain; but see below) and a determiner *ḥetta* ‘(not) even a’ (< Classical Arabic *ḥattā* ‘even, until’):

(129) (a) ʕazīz **ma** taygūl li-ya **wālu**
 Aziz NEG say.IMP.3MSG to-me n.thing
 ‘Aziz doesn’t tell me anything.’ (Adila 1996: 110)

(b) A: mā-l-ək
 what-to-you
 ‘What’s the matter?’

B: wālu wālu
 ‘Nothing, nothing.’
 (Durand 2004: 111)

(130) (a) ana **ma** klīt **ḥetta** ḥāja mən əl-bārəḥ
 I NEG eat.PRF.1SG even thing since yesterday
 ‘I haven’t eaten anything/a thing since yesterday.’

(b) əl-yom **ma** ja **ḥetta** ḥədd
 today NEG come.PRF.3MSG even anyone
 ‘Today no one came.’

⁷¹ The full indefinite system of Moroccan is summarized in Table 5.3 at the end of this section. The discussion here focuses primarily on the problematic n-words.

(c) **ma** kayswa **ḥetta** bəṣla

NEG be worth.IMPF.3MSG even onion

‘It’s not worth a penny.’

(Adila 1996: 111-112)

(d) A: škun kayskən mʕ-ak

who live.IMPF.3MSG with-you

‘Who lives with you?’

B: **ḥetta** wāḥəd

even one

‘No one.’ (Ouali forthcoming: 9)

From these examples we can see that there are two ways of expressing the meaning ‘nothing’ in Moroccan – *wālu* and *ḥetta ḥāja* – both of which must co-occur with negative *ma* in full sentences (yielding only one semantic negation), and both of which are interpreted as negative in non-sentential utterances (i.e. are n-words). Ouali (forthcoming) views both of these as inherently negative. There are important differences between the two expressions, however. Most importantly, while *wālu* always signals the presence of negation, *ḥetta* is common in affirmative sentences, where it broadly retains the original meaning of Classical Arabic *ḥattā* ‘even’:

- (131) **ḥetta** š-šibāni kayḥəbb lə-bnāt
 even the-old man like.IMPF.3MSG the-girls
 ‘Even an old man (still) likes girls.’ (Harrell 1966: 249)

Moreover, its interpretation as negative or affirmative in certain non-sentential utterances is dependent on the context:

- (132) (a) A: ana ḡadi nṣafər š-šħer l-maji
 I FUT travel.IMPF.1SG the-month the-coming
 ‘I’m taking a trip next month.’

B: **ḥətt-ana**

even-I

‘So am I.’

- (b) A: ḥna ma xarjīn-š
 we NEG leave.PTCP.MPL-NEG
 ‘We’re not leaving.’

B: **ḥətt-ana**

even-I

‘Neither am I.’

(Harrell 1966: 249)

Thus we have a number of reasons to view Moroccan *ḥetta* not as a negative quantifier but as a polarity-neutral item that is strongly (but not exclusively)

associated with negative contexts, just as with the other semi-NPIs *abadan*, *lissa* and *baʕad* above (see Tables 5.1 and 5.2).

First, the item that *ḥetta* is derived from is not negative. Second, it co-occurs with sentential negation without triggering a double negation reading. Finally, an identical item is used in a very similar function in the absence of sentential negation and there is no negative reading. The only apparent obstacle to this analysis is the interpretation of examples such as (130)d and (132)b as negative despite the absence of an overt negator. As noted above, Benmamoun (1997, 2000, 2006), who labels *ḥetta* an NPI, offers no explanation for this item's n-word behaviour. We are in a position to offer an explanation, however. Once again, rather than positing ambiguity between two homophonous items whose meanings are virtually identical aside from their polarity, we can analyse *ḥetta* as a polarity-neutral item whose frequent association with negation allows it to be interpreted negatively in certain non-sentential utterances given an appropriate negative context. This analysis is strengthened by the fact that negatively interpreted *ḥetta X* in the absence of sentential negation is restricted to non-sentential utterances such as those given in (130)d and (132)b. By contrast, *wālu* has no such restriction and is found outside of verbal argument positions in full affirmative sentences, where it still means 'nothing':

(133) (a) ḍār ʕayīra ḥsən mən **wālu** (/ *ḥetta ḥāja)

house small better than nothing (/ even thing)

'A small house is better than nothing.'

- (b) *ṣāhib wāḥed ḥsən mən wālu* (/ *ḥetta ḥədd)
 friend one better than nothing (/ even anyone)
 ‘One friend is better than nothing/no one.’

(Hamid Ouali personal communication)

- (134) *rā-ki mənḥūxa ḡala wālu* (/ *ḥetta ḥāja)
 DEM-you.FSG inflated.Fabout nothing (/ even thing)
 ‘You really are conceited about nothing (/for no reason).’
 (<http://www.3iny3ink.com/forum/t193945.html>)

This strongly suggests that in sentences such as (130)a-c, we simply have sentential negation plus non-negative *ḥetta*, giving rise to a single semantic negation as expected. The situation with *wālu* is not so straightforward, however. Given that *wālu* is interpreted as negative both in non-sentential utterances and in affirmative sentences, including comparatives, such as (133)-(134); and given also that it has no homophones or near homophones which are clearly non-negative, it seems we must come to the conclusion that *wālu* is a negative quantifier and that sentences such as (129)a are true instances of negative concord, something I have argued to be lacking in Classical Arabic and Palestinian/Egyptian (with the possible exception of determiner *wala*).

This becomes somewhat clearer from a historical point of view if we consider *wālu*'s probable etymology. I have not seen suggestions for this in print, but it seems likely that it is derived from the negative determiner *wala* that we saw in Palestinian and Egyptian, plus some other element such as *hu(wa)* ‘he, it’ (although the long /ā/ preceding the // is unexpected on this scenario). Although *wala* is not found as a

determiner in Moroccan, it must have been present in earlier varieties of western Maghrebi Arabic given that it has been borrowed into both Tuareg (see (26) in section 2.2) and Northern Berber (as *ula*, Lameen Souag personal communication). If it is correct that inherently (morphologically) negative *wala* is part of the phrase from which *wālu* is derived then it makes sense from a historical perspective that *wālu* itself should also be inherently negative and should also require the presence of sentential negation in negative sentences (rather than affirmative sentences such as (133)-(134)), as this is exactly what we observed for *wala* in (111) and (127). Unlike *wala*, however, since *wālu* has been reanalysed as a single word meaning ‘nothing’, there seems to be no alternative to viewing the single semantic negation that results from its co-occurrence with sentential negation as a case of bona fide negative concord, which we could then analyse in the same manner as suggested for *wala* above.

Thus we find in Moroccan a rather heterogeneous system, where, for example, the concept ‘nothing’ can be expressed by: the negative quantifier *wālu* with the obligatory presence of sentential negation and obligatory absence of *ħetta* (135)a; *ħāja* ‘thing’ obligatorily preceded by *ħetta*, which I have analysed here as a non-negative polarity-neutral indefinite (135)b, or *šay* ‘anything’ (< *šay?* ‘thing’), which is uncontroversially a strong NPI and not an n-word, and may optionally be preceded by *ħetta* (135)c.

- (135) (a) ma šəft (*ħətta) wālu
 (b) ma šəft *(ħətta) ħāja
 (c) ma šəft (ħətta) šay
 NEG see.PRF.1SG even (no)thing

'I didn't see anything.'

(Adila 1996: 109-111)

Further evidence of the heterogeneity of the system comes from the fact that negation marking in sentences containing the weak NPI pseudo-verb *šammər-* 'ever' (cf. Egyptian/Palestinian *šumr-* in Table 5.2) can occur just on *šammər-* (136)a, just on the main verb (136)b, or on both at once (136)c, with precisely the same interpretation (one semantic negation) in each case.

- (136) (a) ma šammər-ni šəft-u
 (b) šammər-ni ma šəft-u
 (c) ma šammər-ni ma šəft-u
 (NEG) ever-me (NEG) see.PRF.1SG-him

'I've never seen him.'

(cf. Caubet 1996: 91, Adila 1996: 105, Durand 2004: 198)

Thus Moroccan appears to wholly resist neat classification as a 'negative concord language' or a 'non-negative concord language', though it has clearly undergone significant developments in this domain with respect to both Classical and Palestinian/Egyptian. Its indefinite system is summarized in Table 5.3 (see Tables 5.1 and 5.2 for legend).

Table 5.3 Indefinites in Moroccan Arabic

	negative quantifier	n-word	strong NPI	weak NPI
Determiner	—	ħetta	→ (ħetta) ^a	(ši) ^b
Thing	wālu ←	wālu	šay	(ši ħāja)
		ħetta ħāja →	(ħetta ħāja)	
Person	—	ħetta ħedd	→ (ħetta ħedd)	(ši) ħedd
		ħetta wāħed →	(ħetta wāħed)	(ši wāħed)
Extent adverb	—	abadən	→ abadən	—
			šay	
			gāf	
Time	ma fammər-	—	—	fammər-
still/yet ^c	—	—	—	—
Conjunction	wala ←	wala	—	—
Place	—	—	layn	(ši maħəll)

^a Like Palestinian/Egyptian *baʕad / lissa*, *ħetta* is polarity-neutral, albeit strongly associated with negation.

^b Phrases with the determiner *ši* are common in ordinary declarative as well as weak NPI contexts.

^c The meaning 'still' is expressed by the auxiliaries *bāqi* and *mazāl*. I have no data on whether these items are n-words or not.

5.2.5 N-words in Maltese⁷²

Maltese appears to be the only Arabic variety that may accurately and straightforwardly be described as a negative concord language. The Maltese indefinite system has been described in some detail by Haspelmath and Caruana

⁷² Recall that in standard Maltese orthography <x> represents /ʃ/.

(1996). They contrast two series, which they call the *xi*-series and the *ebda*-series, respectively (Haspelmath and Caruana 1996: 215):

Table 5.4 Haspelmath and Caruana's (1996)

Maltese *xi*- and *ebda*-series

	<i>ebda</i> -series	<i>xi</i> -series
Determiner	ebda	xi
Thing	xejn	xi ħaġa
Person	ħadd	xi ħadd
Time	qatt	xi darba
Place	imkien	xi mkien

While the members of the *xi*-series are all non-negative and cannot ordinarily occur in the scope of negation,⁷³ the *ebda*-series items are all n-words: when they function as verbal arguments (or predicate-level adjuncts in the case of *qatt* and *imkien*) in a full sentence they must co-occur with the sentential negator *ma* and these sentences are interpreted as having a single semantic negation; however, they can appear without *ma* in non-sentential utterances which are nevertheless still interpreted as negative, as can be seen in (137) (= Haspelmath and Caruana's (16a) and (17)).

- (137) (a) It-tifla **ma** rat **xejn**.
 the-girl NEG see.PRF.3FSG n.thing
 'The girl didn't see anything.'

⁷³ Example (141) below clearly contradicts the second half of this statement. A more accurate generalization appears to be that members of the *xi*-series can only appear in the scope of negation when used predicatively in negative copular sentences.

(b) A: X'rat?

what-see.PRF.3FSG

'What did she see?'

B: **Xejn!**

n.thing

'Nothing!'

(Maltese)

Given our above observations concerning n-words and negative concord, however, it seems too hasty of Haspelmath and Caruana (1996: 217) simply to cite (137)b (their (17)) as proof of the inherently negative status of all members of the *ebda*-series. This is particularly clear, in fact, from the fact that the n-word *qatt* '(n)ever', unlike the other members of the *ebda*-series, is regularly found with non-negative meaning in nonveridical contexts such as questions and conditionals, as in (138) (= Haspelmath and Caruana's (24)), showing that *qatt* at least cannot be a single, inherently negative lexical item.

(138) (a) Jekk **qatt** tiġi Londra, ejja arani.
 if ever come.IMPF.2s London come.IMP.2S see.IMP.2S-me
 'If you ever come to London, come and see me.'

(b) **Qatt** mort Londra?
 ever go.PRF.2s London
 'Have you ever been to London?' (Maltese)

Similarly, *imkien* ‘n.where’, while much more common in negative sentences, can also be used in affirmative declarative contexts, simply meaning ‘place’ (although the plural *imkeyyen* ‘places’ is far more common than the singular, which must be regarded as an archaism in such a context; Michael Spagnol personal communication). The same is not true, however, of the determiner *ebda*, the ‘person’-pronoun *ħadd* or the ‘thing’-pronoun *xejn*, about which there cannot be much doubt that they are inherently negative. This is perhaps clearest for *xejn*, which has given rise, by a regular process of gemination of the second root segment, to a denominal verb *xejjen* ‘to make nothing, destroy’:

(139) *lżda xejjen lil-u n-nifs-u.*

but nullify.PRF.3MSG to-him the-self-his

‘But he made himself nothing.’ (Philippians 2: 7)

Moreover, like Moroccan *wālu* and Palestinian/Egyptian *wala*, *xejn* can appear in non-verbal-argument positions in affirmative sentences, including comparatives, where it retains its negative meaning:

(140) *Issa noffru KMiles b’xejn!*

now offer.IMP.1PL KMiles for-nothing

‘Now we’re offering KMiles for free!’ (HSBC/Air Malta advertisement, Malta)

- (141) Dan m'huwiex xi kumpens enormi imma dejjem huwa
 this NEG-it-NEG some compensation enormous but always it
 aħjar minn **xejn**.
 better than nothing
 'This isn't an enormous amount of compensation, but still it's better than
 nothing.'
 (www.parliament.gov.mt/file.aspx?f=1783)

The same is true for *ħadd* and *ebda*:

- (142) Dik ħajjitha u kuxjenza tagħha u ta' **ħadd** iżjed
 that life-her and conscience of-her and of no.one more
 'That's her life and her conscience and no one else's.'
 (<http://forum.huntinginmalta.org.mt/YaBB.pl?num=1218971916/3>)
- (143) Drapp rożż ta' **ebda** ħtieġa li jinħasel.
 material rough of no need COMP be.washed.IMPF.3MSG
 'Rough material with no need for washing.'
 (<http://www.sanpawlmalta.com/Fratellanza/index.htm>)

Thus we have robust evidence that *ebda* 'no', *xejn* 'nothing' and *ħadd* 'no one' are negative quantifiers. *Qatt* 'ever', on the other hand, definitely is not and *imkien* possibly should not be analysed as such given its residual function as a noun simply meaning 'place(s)'. Thus Haspelmath and Caruana's *ebda*-series items are similar in one respect (they are all n-words), but differ somewhat in their behaviour. There are

a couple of apparent difficulties even with the analysis of *ebda*, *xejn* and *ħadd* as negative quantifiers, however. Haspelmath and Caruana (1996: 219) state that ‘*ebda*-indefinites may also occur in the standard of comparison, or in the qualifying relative clause of a superlative expression’. The example they give – (144) (= their (23)) – features *qatt* and *ħadd*.

- (144) Dan huwa l-isbaħ inkwatra li ħadd qatt pinga
 this it the-beautiful.CMPR picture COMP n.one ever paint.PRF.3MSG
 ‘This is the most beautiful painting that anyone has ever painted.’

Native speakers I have consulted accept this sentence as grammatical, although some describe it as only marginally acceptable. All agree that without *qatt* the sentence is highly marginal, and uncontroversially non-negative *xi ħadd* ‘anyone’ is required to rescue grammaticality, *xi ħadd* also being preferred even when *qatt* is present. *Xejn* and *ebda* appear never to be possible in the standard of comparison (unless with their ordinary negative meaning as in (141)), and members of the *xi*-series are very common, as in the following example:

- (145) L'idea taż-żewġ awturi... certament aħjar minn xi omeliji li
 the-idea of-the-two authors certainly better than some sermons COMP
 smajt
 hear.PRF.1SG
 'The idea of those two authors... is certainly better than some sermons I've
 heard.' (<http://www.timesofmalta.com/articles/view/20090625/local/church-film-reviewing-board-chairman-removed-from-post>)

The use of non-negative *ħadd* in cases such as (144) thus appears to be a marginal relic of a former less restricted distribution, which does not substantially affect the analysis of this item (and certainly not that of *xejn* and *ebda*) as a negative quantifier. Unsurprisingly, there are also fixed relic expressions with non-negative meanings containing these items, such as *qabelxejn* 'first of all' (*qabel* 'before'), *kull xejn* 'everything' (*kull* 'all'), *kulħadd* 'everyone' and *ħadd ieħor* 'someone else' (*ieħor* 'other'; compare *ħadd iżjed* 'no one else' (*iżjed* 'more')). These too should be interpreted as fossilized pointers towards the meanings that *xejn* and *ħadd* once had more generally, rather than genuine manifestations of their enduring non-negativity.

This naturally raises the question of how the present Maltese indefinite system evolved from one which we must assume originally closely resembled that of Classical Arabic and the north African dialects. Let us first consider the etymology of the individual items. The determiner *xi* is essentially identical in meaning, form and function to Moroccan *šī* (see Table 5.3). This item is evidently derived from Classical Arabic *šay?* 'thing', presumably through reanalysis of a genitive construct with partitive meaning: 'a thing of *X*' > 'some/any *X*'. The items *ħaġa*, *darba* and *(i)mkien*,

with which *xi* frequently co-occurs, are lexical items (meaning ‘thing’, ‘time’ (< ‘path’) and ‘place’, respectively) in their own right, so it is unclear whether these collocations should be seen as fixed expressions or simply specific instances of nouns determined by *xi*.⁷⁴ In any case, their etymology is trivial. *ħadd* is also quite clearly the same item etymologically as Moroccan *ħədd* and Classical *aħad* ‘anyone’. As we have seen, however, *ħadd* has undergone a marked contraction in the environments in which it can occur: from all weak NPI contexts as expected by comparison with *ħədd* and *aħad*, to essentially only negative contexts, culminating in its becoming a negative quantifier, except in certain fossilized phrases including *xi ħadd* ‘someone/anyone’. *Qatt* ‘ever’ is clearly derived from Classical *qaṭṭu* ‘ever’ (see Table 5.1), which is interesting in itself, given that this item is not found in other north African Arabic dialects. Also, the fact that *qatt* is an ordinary weak NPI, frequently found in non-negative nonveridical contexts, is interesting given that its ancestor *qaṭṭu* appears to have been a strong NPI. Either we have here a rare instance of a polarity item becoming weaker over time, or, perhaps more likely, the restriction of *qaṭṭu* virtually only to negative contexts could have been a development of immediately post-Islamic written Arabic (i.e., what became Classical Arabic), not shared by the contemporaneous spoken varieties.

The remaining n-words *ebda* and *xejn* have undergone a somewhat greater shift. *Xejn* is clearly also derived from *šayʔ* ‘thing’, apparently with a vestige of the indefinite accusative suffix *-an*. A number of offshoots of *šayʔ* have developed into a

⁷⁴ Although, as Michael Spagnol pointed out to me, English words such as *something* and, probably also, *sometime/someday*, which would be likely to be analysed as lexicalized, fixed expressions, seem to be direct parallels to Maltese *xi haġa* ‘something’ and *xi darba* ‘sometime(s)’.

range of weak and strong NPIs in a range of Arabic dialects (cf. example (58) in section 2.4), but it is only in Maltese that one of these has become a negative quantifier. *Ebda* appears to be derived from *abadan* ‘ever’, and, if this is correct, this would make Maltese also the only Arabic variety in which this item has become a determiner. This development from indefinite ‘time’-adverb to negative determiner does not seem to be common in other languages, but it is attested in English, for example, in the colloquial or archaic forms *nary (a)/ne’er (a)* ‘no *X*’ < *never*.

Concerning the development of negative concord, we have seen that while there are negative concord tendencies in Moroccan, Maltese appears to be the only Arabic variety with undoubted negative quantifiers in the categories determiner, ‘thing’-pronoun and ‘person’-pronoun. Moreover, when either of these types of pronouns or a determiner is required in the scope of negation, only the negative quantifier versions are grammatical. This situation is not found in other Arabic varieties: Maltese is the only one with obligatory negative concord. In accounting for this unique position, it is tempting to point to the centuries of intensive contact between Maltese and the Romance varieties of Sicily, which have always exhibited negative concord since their earliest attestations (Adam Ledgeway personal communication), as in the following example from the 14th century:

(146) dicendu que **nullu** homu **non** ci era trassutu
 saying COMP no man NEG there be.PST.3SG enter.PTCP
 ‘saying that no man had entered there.’

(C14th Sicilian; Parry 2009)

In an indefinite system in which no indefinites are marked negative morphologically, negative sentences containing indefinite pronouns will always be analysable either as featuring negative concord between a marker of sentential negation and a negative quantifier, or as featuring a non-negative pronoun and therefore no negative concord. Maltese speakers, like speakers of other Arabic varieties, will originally have made the latter analysis. However, intensive L2 exposure to Sicilian could have prompted a reanalysis of the polarity of the indefinites of their native Maltese on the basis of the morphologically transparent negative concord system of Sicilian. This would amount to borrowing (under RL agentivity) of the rule that the presence of sentential negation requires indefinite pronouns in the scope of negation to be negative quantifiers. Alternatively, and perhaps more plausibly, this change could have occurred under SL agentivity when L1 Sicilian speakers (for example in mixed Sicilian-Maltese marriages) identified the non-negative indefinite pronouns of their L2 Maltese with the negative indefinite pronouns of their L1 Sicilian and therefore imposed the semantic feature of negativity onto these items, including in affirmative contexts such as (140)-(143). On the other hand, given that negative concord is rather common in the world's languages (Kahrel 1996), it may be that the development of negative concord in Maltese was a purely internal process that was unaffected (or perhaps merely accelerated) by contact with Sicilian.

The synchrony of the Maltese indefinite system is summarized in Table 5.5 and its evolution in Table 5.6 (arrows in Table 5.5 represent analyses of n-words as in previous tables; in Table 5.6 arrows indicate diachronic developments from the presumed inherited to the present-day state of affairs).

Table 5.5 Indefinites in Maltese

	negative quantifier	n-word	strong NPI	weak NPI
Determiner	ebda	← ebda	—	(xi) ^a
Thing	xejn	← xejn	—	(xi ħaġa)
Person	ħadd	← ħadd	—	(xi ħadd)
Extent adverb	lanqas	xejn ← lanqas	—	—
Time	—	qatt	→	qatt (xi darba)
still/yet ^b	—	—	—	—
Conjunction	u lanqas	← u lanqas	—	—
Place	—	imkien	→	(imkien) ^c (xi mkien)

^a As noted above, phrases containing *xi* (like the very similar Moroccan *šī*) occur in ordinary declarative as well as weak NPI contexts.

^b 'Still' is expressed by the auxiliary *għad* in Maltese. It is polarity-neutral and not an n-word.

^c Like Palestinian/Egyptian *baʕad / lissa* and Moroccan *ħetta*, *imkien* is polarity-neutral, albeit strongly associated with negation (a semi-NPI).

Table 5.6 The evolution of Maltese indefinites

		negative quantifier	NPI	polarity-neutral item
Determiner	inherited	—		ši
	present-day	ebda		xi
Thing	inherited	—		šayʔan, ɥāja
	present-day	xejn		xi ɥaɣa
Person	inherited	—	ɥadd	wāɥid
	present-day	ɥadd		xi ɥadd
				wieɥed
Time	inherited	—	abadan, qaɥtu	—
	present-day	—	qatt	xi darba
Place	inherited	—	—	makān
	present-day	—		imkien

5.3 Stage III negation in Palestinian (and Cairene) Arabic

Having dealt with the genesis of stage II negation and developments in indefinites, the final section of this chapter looks at the progression to stage III negation in certain Arabic varieties. As mentioned in section 1.3.1, a stage III construction is attested in a number of dialects, including Palestinian, Šaʔdī and Omani. In none of these varieties, however, has the syntax of stage III negation been studied in any detail before now. The focus of this section is on stage III negation in Palestinian, illustrated with a popular bilingual pun in (147).

(147) fi-š samak kol faḥm
 there.is-NEG fish eat.IMP.2MSG coal

‘If there’s no fish, eat coal.’

(Palestinian)

Building on fieldwork I carried out in Israel and the Occupied Palestinian Territories in the spring of 2008 to investigate this phenomenon, in this section I propose a reconstruction of how the stage III construction originally came about in Palestinian as well as to a limited extent, and rather less obviously, also in Cairene.

As will become clear in what follows, there are a number of rather idiosyncratic-looking restrictions on the distribution of stage III negation in Palestinian, while in Cairene the contexts in which the formally similar construction is possible are sharply circumscribed, and appear to be wholly orthogonal to those relevant for Palestinian.

The nature of these two sets of restrictions leads to two separate reconstructions of how purely postverbal negation arose in the two dialects. For Palestinian, I argue that, despite initial appearances, a synchronic rule that deletes *mā* in a particular phonological environment cannot account for the data. Instead, I reconstruct an earlier, more constrained process of phonological reduction towards zero, which then triggers a syntactic reanalysis – making stage III negation a syntactic reality – leading to the subsequent spread of this construction to related syntactic contexts. For Cairene, I will argue that postverbal negation develops not from the loss of *mā* in negative sentences, but from the gradual negativization of *-š* in contexts where it has long appeared with non-negative meaning: specifically, the weak NPI contexts of

interrogatives and conditionals. It will be seen that this development has a direct parallel in the history of French.

The rest of the section is structured as follows. Section 5.3.1 gives a summary of the sparse data on stage III negation in Palestinian and Cairene in the previous literature. Section 5.3.2 presents the research questions that are prompted by this existing data, while sections 5.3.3 and 5.3.4 describe the methodology and the findings of the fieldwork carried out to provide answers to some of these questions. In section 5.3.5 I propose separate reconstructions for postverbal negation in Cairene and Palestinian, and in section 5.3.6 I make some tentative predictions concerning future developments.

5.3.1 Previous literature

The issue of the loss of *mā* in Palestinian appears not to have been the subject of detailed discussion in previous literature, though it has often been commented on in passing. For example, Driver (1925: 197) notes that *mā* is “occasionally” dropped in negative sentences, while Bauer (1913: 122) states more specifically that “die Verneinung des [Imperfekts] wird manchmal nur durch suffigiertes *š* ausgedrückt.”⁷⁵ The same view, albeit with reversed emphasis, is expressed in later studies, for example Obler (1975: 101, 105): “Suffix *-š* is the general verbal negator... Perfect verb forms regularly add the *mā* to their negated Vs,” and Shahin (2000: 37): “A V is negated with *-š*, for the perfect *mā* is also used.” On the basis of this apparent tense split Blau (1960: 198) hypothesizes that *mā* is deleted in virtue of its phonological

⁷⁵ “The negation of the imperfect is sometimes expressed by suffixed *š* alone.”

similarity to the *b*-prefix that is characteristic of the imperfect indicative in Levantine and Egyptian Arabic.⁷⁶

A recent, and strongly expressed, version of this idea comes from Hoyt (2006: 16) who states: “Omission of *mā-* is possible only with stems beginning with labial obstruents [b] or [f].” On the face of it, this statement is contradicted, however, by Obler’s (1975: 107) observation that the pseudo-verbs,⁷⁷ not all of which begin with [b] or [f], “seem to be in free variation with respect to the optionality of *mā-*.” It is not clear that this last observation is entirely accurate, however. Obler gives a number of examples from her corpus, but only with *fi* ‘there is’ and *bidd-* ‘to want’ does she provide examples both with and without *mā-*. In particular, there are in fact no examples in her corpus of *ʕind-* or *il-* (both meaning ‘to have’) negated with the stage III construction (i.e. without *mā-*). We will see from the results presented in section 5.3.4 that the stage III construction is indeed ungrammatical with *ʕind-*, as well as with the perfect of regular verbs, and that these data can provide the basis for a principled reconstruction of the changes that gave rise to postverbal negation in Palestinian.

A related issue that has often been noted in passing in grammars, but has never to my knowledge been discussed in detail, is the phenomenon of an enclitic *-š*

⁷⁶ This verb form is often referred to as the *b*-imperfect. This contrasts with the unprefixated imperfect, sometimes called the *y*-imperfect, which tends to have a subjunctive or irrealis function in these varieties.

⁷⁷ These are a small, closed-class set of highly irregular verbs derived mainly from prepositional phrases, whose person inflections are identical to the oblique pronominal suffixes from which they derive. The commonest pseudo-verbs are *fi* ‘there is’, *ʕind-/ʕand-* ‘to have’, *maʕ-* ‘to have (on one’s person)’ and *bidd-* ‘to want’. *il-* is a less commonly used pseudo-verb (at least in the dialects focussed on here) also meaning ‘to have’ (see footnote 82).

apparently marking questions (some authors claim only negative questions) in various dialects, possibly including Palestinian. For example, Driver (1925: 197) states (of Palestinian) that “the enclitic *-š* often stands alone, without a preceding *mā*, in negative questions.” Note his use of “often” in contrast to the “occasionally”, cited earlier, with which he describes the occurrence of the stage III construction in negative declarative contexts. He gives the following example with, note, a perfect verb:

- (148) *nattilt-iš* *?ijray-k*
 wash.PRF.2MSG-NEG foot.DU-your
 ‘Haven’t you washed your feet?’ (Palestinian; Driver 1925: 197)

More recent studies of Palestinian Arabic, however, such as Shahin (2000) and Opler (1975: ch.4) (which focuses specifically on reflexes of *šay?* in a rural Palestinian dialect), make no mention of any discrepancy between negation strategies in declarative versus interrogative clauses.

This contrasts with the situation in various north African dialects (here I will focus on Cairene), which do not allow negation of declarative clauses with a stage III construction (i.e. they require *mā* in addition to *-š* in negative declarative clauses), but do feature the question-marking enclitic *-š* (without *mā*), which does not appear to exhibit the tense restrictions reported for the postverbal negative construction in Palestinian. There is disagreement in the various grammars that mention this construction as to whether it encodes negation in addition to interrogation, or whether in fact preverbal *mā* is required in all cases of negation, be they declarative or

interrogative. For example, the early grammars of Spitta-Bey (1880) and Willmore (1901) analyse this construction as inherently negative, as in the following:

- (149) *kunt-iš* *hināk*
 be.PRF.2MSG-NEG.Q there
 ‘Weren’t you there?’ (Egyptian; Willmore 1901: 298)
- (150) *ʕandak-š*i** *qiršēn*
 have.2MSG-NEG.Q penny.dual
 ‘Don’t you have two pennies?’ (Egyptian; Spitta-Bey 1880: 416)

Gamal-Eldin (1967) offers the same analysis. However, various others, including, interestingly, Gary and Gamal-Eldin (1981), state that this construction simply marks questions, and some expressly contradict the idea that it also involves negation. Davies (1981: 278), for example, refers to the analysis of Spitta-Bey and Willmore as “confusion”.

In reality, to judge from the examples in Woidich (1968, 2006), as well as some of those given by Willmore (1901), any analysis of enclitic *-š* without preverbal *ma* in Cairene which sees it as unambiguously affirmative on the one hand, or unambiguously negative on the other, must be wrong. In (151), for instance, the context makes clear that *-š* cannot be negative: the speaker must be asking whether he is in the wrong street, not whether he has failed to get the wrong street:

- (151) bēt abū-ya huwwa fēn walla akun-šī
 house father-my it where or be.IMP.F.SBJV.1SG-Q
 ġliṭṭi fi š-šāriṯ
 err.PRF.1SG in the street
 'Where's my father's house? Or have I got the wrong street?'
 (Woidich 2006: 358)

In (152) on the other hand, in which there is an embedded interrogative clause, the second -š must be negative for the predicate it attaches to to make sense as an alternative to the option that precedes it (the first -š is uncontroversially part of a negative construction).

- (152) ma ṯraf-šī kān mawgūd wala kān-šī
 NEG know.IMP.F.SBJV.1SG-NEG be.IMP.F.3SG present or be.PRF.3SG-NEG.Q
 'I don't know if he was present or not.' (Willmore 1901: 298)

Consider also (153), which, given the information that the speaker did in fact give up growing cotton and plant rice, is unambiguously negative.

- (153) law kunt-iš sibṯ il-ṯuṯni w-zaraṯi
 if be.PRF.1SG-NEG.? abandon.PRF.1SG the-cotton and-plant.PRF.1SG
 ruzz kān...
 rice be.PRF.3SG
 'If I hadn't given up cotton and grown rice, then...' (Woidich 1968: 107)

Note also that (153) involves a conditional rather than a question.⁷⁸ As such *-š* without *ma* in Cairene in fact seems to have the distribution of a weak NPI, with a function similar to English *at all* in (154)-(155), rather than being specifically a question marker.⁷⁹

(154) Can I interest you in our brochure at all?

(155) If you see John at all, tell him I called.

However, it appears that *-š* without *ma* in present-day Cairene must express negation when it appears in conditionals, as in (153), in contrast to the situation in questions where it can be either negative or affirmative, as shown in (151)-(152). Willmore (1901) and Woidich (1968, 2006) give several examples of negative conditionals featuring *-š* without *ma* but no affirmative ones. Similarly, a search of Egyptian internet chat rooms turns up many examples of clearly negative conditionals featuring *-š* without *ma*, but none that are clearly affirmative. Native-speaker linguists that I have consulted on this point also reject an affirmative interpretation of *-š* without *ma* in conditionals in Cairene Arabic. So there does appear to be a clear difference in the behaviour of this item in interrogatives (optionally negative) versus conditionals (obligatorily negative). Clearly, where *-š* without *ma* is negative, it is no longer parallel to English *at all*, which cannot on its

⁷⁸ Willmore (1901: 298) gives a similar example.

⁷⁹ Note in this connection that *-š* is never obligatory in questions, and lends them, according to Woidich (2006: 306) a “dubitativen Sinn” (“a doubtful sense”), which appears also to be the case with English *at all*.

own identify negation in questions and conditionals. The facts concerning *-š* in Cairene are shown in Tables 5.7 and 5.8.

Table 5.7 Distribution of *-š* without *ma* in Cairene

Context	Grammaticality of <i>-š</i> without <i>ma</i> with a negative interpretation	Grammaticality of <i>-š</i> without <i>ma</i> with an affirmative interpretation
declarative clauses	*	*
interrogative clauses	✓	✓
conditional clauses	✓	*

Table 5.8 Grammaticality of negative constructions by context in Cairene

Context	Bipartite construction	Postverbal construction
declarative clauses	✓	*
interrogative clauses	✓	✓
conditional clauses	✓	✓

To summarize: the presence of *-š* without *ma* in Cairene is restricted to conditional and interrogative clauses. Where it is found in interrogative clauses it can either express negation or merely lend the question a “doubtful sense”, whereas in conditionals it can only express negation.

5.3.2 Research questions

A number of questions, both empirical and theoretical, arise from the above data. Most importantly for present purposes: how can we explain the development of stage

III negation in Palestinian Arabic, and is the question-marking or NPI *-š* found in Cairene and other dialects part of the answer? Also to be considered in this connection is the apparent tense split in Palestinian negation: is it merely a statistical tendency that *mā* is dropped more often with the imperfect than with the perfect, or is it in fact obligatory with the perfect, and what, if anything, can this tell us about the origins of the postverbal construction? And a final closely connected question: can the phonological deletion rule suggested by Blau (1960), and explicitly stated by Hoyt (2006), do the job of explaining the data concerning the distribution of bipartite and postverbal negative constructions both between tenses and between regular and pseudo-verbs?

5.3.3 Fieldwork methodology

In order to obtain some answers to these and related questions I compiled an acceptability judgment questionnaire which was administered to 29 sedentary (i.e. non-Bedouin)⁸⁰ native speakers of Palestinian Arabic resident in a range of locations in Israel, the West Bank and the Golan Heights. Approximately two thirds of respondents were male, the rest female, and they ranged from teenagers to individuals in their sixties.

The questionnaire, which is reproduced with glosses and translations in the appendix at the end of this thesis, comprised 33 sentences. For each of these informants were asked to give an acceptability judgment ranging between three predetermined values: *šādī* 'normal, fine', *yarīb* 'odd' and *mustaḥīl* 'impossible'. It was

⁸⁰ Recall that traditional Bedouin dialects lack both the stage II and stage III constructions.

made clear that judgments should be made on the form of the sentences rather than their meaning. The following is a sample sentence in transliteration:

- (156) (ana) akalt-iš il-fūl
 I eat.PRF.1SG-NEG the-fava beans
 'I didn't eat the fava beans.'
 ʕādī
 ʕarīb
 mustahīl

There were two reasons for choosing an acceptability judgment questionnaire as the main research tool despite the non-naturalistic data that this kind of a questionnaire provides (Schütze 1996, Cowart 1997). Firstly, some of the phenomena I was investigating are sufficiently infrequent that they can only be relied on to occur a useful number of times in a corpus of spoken language far larger than would have been practical to record, transcribe and analyse for the purposes of this research. Secondly, even if such a corpus were available, it would not furnish the kind of negative evidence – on what strings are ungrammatical in Palestinian Arabic – that can be obtained from acceptability judgments, and thus could not answer some of the key research questions outlined above.

5.3.4 Results and discussion

93% (27/29) of informants judged acceptable the stage III construction with an imperfect verb such as *baḥibb* ‘I like’, as shown in Table 5.9:⁸¹

Table 5.9 Acceptability judgments for stage III construction with imperfect

Form	Percentage judged fine	Percentage judged odd	Percentage judged impossible
<i>baḥibb-iš</i>	93 (27)	7 (2)	0 (0)

It is surprising, perhaps, that two informants judged this construction odd, given that it is uncontroversially a basic feature of the grammar of Palestinian Arabic. Although it is possible that they were objecting to something other than the grammaticality of the relevant structure (exemplified in (157)), these two informants are excluded from consideration in the presentation of results for all remaining questions that assume the grammaticality of the stage III construction with an imperfect verb.

- (157) (ana) baḥibb-iš il-fūl
 I like.IMPF.1SG-NEG the-fava.beans
 ‘I don’t like fava beans.’

⁸¹ All the cells in Tables 5.9-5.14 show the percentage value first, followed by the absolute figure. Also, all verbs in these tables are cited in first person singular (or plural in the case of Table 5.13) forms apart from the impersonal existential *fī*.

Table 5.10 shows the responses of the 27 informants who judged (157) acceptable to sentences such as (156), that is, sentences in which a perfect verb is negated with the stage III construction, and also to sentences where a perfect verb is negated with the stage II construction.

Table 5.10 Acceptability judgments for stage II/III constructions with the perfect

Form	Percentage judged fine
<i>mā akalt-ʾš</i>	100 (27)
<i>akalt-ʾš</i>	4 (1)

The rather striking result is that while every one of these 27 informants judged the stage II construction with the perfect to be fine, just one person judged the stage III construction with the perfect to be fine, and in fact this individual volunteered the information that her response was influenced by the kinds of structures that she recognised from the speech of her young children, a point to which we return in section 5.3.6.

This result strongly substantiates the statements of the authors in section 5.3.1 to the effect that the postverbal construction is only found with the imperfect and not the perfect. But what of the related claim of Hoyt's (2006: 16) that: "omission of *mā* is possible only with stems beginning with labial obstruents [b] or [f]"? In view of the discrepancy illustrated in Tables 5.9 and 5.10 between the acceptability of the stage III construction with the *b*-imperfect versus its unacceptability with perfect verbs

beginning with other sounds, this claim does have an initial plausibility. Indeed, this plausibility is increased if we take a critical look at Obler's (1975: 107) above-mentioned observation that the pseudo-verbs "seem to be in free variation with respect to the optionality of *mā*." Several sentences in the questionnaire aimed to assess the accuracy of this observation. The responses to these sentences (summarized in Tables 5.11 and 5.12) broadly substantiate Hoyt's statement and directly contradict Obler's in its current general form, in that while the postverbal construction was very widely accepted by informants for the pseudo-verbs *fi* 'there is', *maʕ*- 'to have (on one's person)' and *bidd*- 'to want' (Table 5.11), it was almost universally rejected for the pseudo-verb *ʕind*- 'to have' (Table 5.12).⁸²

Table 5.11 Acceptability judgments for stage II/III constructions with labial-initial pseudo-verbs

Form	Percentage judged fine
<i>mā biddī-š</i>	
<i>biddī-š</i>	100 (27)
<i>mā maʕī-š</i>	
<i>maʕī-š</i>	
<i>mā fi-š</i>	96 (26)
<i>fi-š</i>	

⁸² In a pilot version of this questionnaire I included several sentences designed to test the acceptability of the postverbal construction also with the pseudo-verb *i/-* 'to have'. I have not found any native speaker of Palestinian Arabic who judges a string such as *(i)lī-š* 'he does not have' acceptable. However, I do not present results here because, in fact, it proved difficult to find any sentence at all containing the pseudo-verb *i/-* (including affirmative sentences) that was judged acceptable by a clear majority of informants. This seems to be a reflection of the ongoing obsolescence of *i/-* as an expression of verbal possession in Palestinian Arabic, as noted by Rosenhouse (2003/4).

Table 5.12 Acceptability judgments for stage II/III constructions with pseudo-verb *ʃind-*

Form	Percentage judged fine	Percentage judged odd	Percentage judged impossible
<i>ʃindī-š</i>	4 (1)	48 (13)	48 (13)
<i>mā ʃindī-š</i>	70 (19)	30 (8)	0 (0)

Table 5.11 shows that there is near-universal acceptance of the stage II construction with the three pseudo-verbs with initial labials: *fī*, *maʃ-* and *bidd-*. This is in keeping with the spirit of Hoyt's statement that *mā* can only be omitted before the labial obstruents [b] and [f], though it would be more accurate, restricting ourselves to the data concerning just the pseudo-verbs for the moment, to say that *mā* can be omitted before any labial consonant.

Table 5.12 shows a very different story for the pharyngeal-initial *ʃind-*. First of all there is the somewhat surprising result that only 70% of speakers accepted a sentence in which *ʃind-* was negated with the stage II construction. To judge from supplementary information given by many of the 30% who considered this structure odd, the acceptance rates here would appear to be depressed by a sense that an alternative periphrastic construction *fī-š ʃind-*, in which *ʃind-* is a preposition 'to/for/at' rather than a verb, and the existential *fī* carries the negation, is more common.⁸³ It may also be significant that these 30% judged *mā ʃindī-š* 'I do not have' merely odd,

⁸³ Perhaps because of the greater agentivity and alienability associated with the possessive meaning of *maʃ-* 'to have (on one's person)' there does not appear to be a similarly frequent parallel construction *fī-š maʃ-*.

rather than impossible. However, we should probably be cautious about reading too much into the figures for strings judged odd versus impossible presented in any of the tables in this section. I deliberately left the distinction between odd and impossible strings undefined for the purposes of this questionnaire, and only made the distinction in order to allow informants a middle way between total acceptance and total non-acceptance. My impression while administering the questionnaire was that informants tended to make a binary choice for each sentence between fine or not, and if it was not fine then they appeared to choose a value of odd or impossible either at random or based on factors unconnected with the level of grammaticality of the string in question.

In any case, it is clear from Table 5.12 that almost no informants judged the postverbal construction with *šind-* to be acceptable. As such, still restricting ourselves just to the data on the pseudo-verbs, we have clear support for the suggestion that *mā* can only be omitted before a verb beginning with a labial. Note also that the issue of a tense split does not arise here, since the pseudo-verbs lack a non-periphrastic past tense, a point which will be important to the reconstruction of the development of the postverbal construction proposed below.

If, however, we extend our purview beyond the pseudo-verbs, the connection between the postverbal construction and phonology disappears. Opler (1975: 109) already points out one sense in which the presence of a following labial cannot, generally speaking, be a necessary condition for the omission of *mā*: a very common context for the postverbal construction is that of prohibitives as in (158). These are formed by negating the *y*-imperfect in the second person, which always begins with the alveolar stop [t] (not [b]).

(158) txafi-š

fear.IMP.SBJV.2MSG-NEG

‘Don’t be afraid.’

In fact, as can be seen from Table 5.13, the presence of a following labial does not appear to be a sufficient condition for omission of *mā* either.

Table 5.13 Acceptability judgments for stage II/
III constructions with a labial-initial perfect verb

Form	Percentage judged fine
<i>mā mesaḥnā-š</i>	100 (27)
<i>mesaḥnā-š</i>	4 (1)

Here we see an identical pattern of responses to the sentence type shown in (159) (i.e. a stage III construction with a labial-initial perfect verb) as was observed for the sentence type shown in (156) (i.e. a stage III construction with a non-labial-initial perfect verb).

(159) mesaḥnā-š iṭ-ṭāwila

wipe.PRF.1PL-NEG the-table

‘We didn’t wipe the table.’

question. It would have been interesting to find that either or both interpretations were possible. In the event, I found that neither is: the sentence in (160) was universally rejected, as can be seen from Table 5.14.

Table 5.14 Acceptability judgments for question-marking *-š* with a perfect verb

Form	Percentage judged fine	Percentage judged odd	Percentage judged impossible
<i>akalt-š...?</i>	0	48 (14)	52 (15)

As such, we can say with some confidence that the question-marking/NPI *-š* found in Cairene and other dialects is not found in present-day Palestinian.⁸⁴ The pre-history of both the Palestinian postverbal negative construction and question-marking/NPI *-š* in Cairene are the topic of section 5.3.5.

The findings described in this section are summarized in Table 5.15.

⁸⁴ Concerning Driver's (1925: 197) example (reproduced in (148)) of a postverbal construction with a perfect verb in an interrogative sentence, two possibilities, which I will not choose between here, suggest themselves: either Palestinian used to have NPI *-š* just as in Cairene in addition to the standard bipartite/postverbal negative construction, but has now lost the former, or Driver was mistaken in citing (148) as a possible sentence in Palestinian. Note in this connection that Bauer's (1913) grammar of Palestinian gives no examples of a postverbal negative construction with a perfect verb.

Table 5.15 Grammaticality of negative constructions by context in Palestinian

	Context	Grammaticality of stage II construction	Grammaticality of stage III construction
Regular verbs	<i>b</i> -imperfect	✓	✓
	prohibitive	✓	✓
	perfect	✓	*
Pseudo-verbs	<i>fi</i>	✓	✓
	<i>maʕ-</i>	✓	✓
	<i>bidd-</i>	✓	✓
	<i>ʕind-</i>	✓	*

While the stage II construction is grammatical with all pseudo-verbs and regular verbs in both the imperfect and perfect tenses, the contexts in which the stage III construction is grammatical are more limited. It is very widely accepted by informants with both the *b*-imperfect and the prohibitive, which is formed by adding the second person *t*-prefix to the imperfect verb, but which lacks the indicative-marking *b*-prefix. However, it is not accepted with any regular verb in the perfect. Among the pseudo-verbs, which lack a non-periphrastic past form, the postverbal construction is very widely accepted with the verbs *fi* 'there is', *maʕ-* 'to have (on one's person)' and *bidd-* 'to want', however it is rejected with *ʕind-* 'to have'.

5.3.5 Reconstruction

The data presented in the previous sections are idiosyncratic to say the least, and would seem to resist explanation in terms of a straightforward synchronic analysis: we have seen for Palestinian that a rule of *mā*-deletion under conditions of phonological similarity both overgenerates (in allowing postverbal negation with

labial-initial perfects as in (159)) and undergenerates (in disallowing postverbal negation with alveolar-initial prohibitives as in (158)). At the same time, a reduction of these data to more basic, independently motivated synchronic syntactic principles seems equally unfeasible.

In Cairene we have a similarly idiosyncratic set of restrictions: enclitic *-š* is generally not capable of expressing negation in the absence of preverbal *mā*, except in interrogative and conditional clauses; where it appears in interrogative clauses it is optionally negative, in conditional clauses it is obligatorily so. Here too, a simple explanation of the data in terms of synchronic syntax, semantics or phonology seems an unlikely prospect. As such, the facts presented would seem particularly amenable instead to an explanation in terms of internal reconstruction. As Givón (2000: 114) observes: “Much like the abductive reasoning practiced by the evolutionary biologist, the [internal reconstruction] practitioner in linguistics takes it for granted that synchronic irregularities are merely the foot-prints of diachronic change from earlier regularities.” The aim of this section is to reconstruct a series of plausible, natural changes that led from an earlier, more regular negation system (i.e. one in which the bipartite construction has its current broad distribution, while the postverbal construction is not possible in any context) to the irregular, idiosyncratic systems we find in Cairene and Palestinian today. Of course, in performing this reconstruction, it will also be essential to consider relevant data from related Arabic varieties.

Thus, the existence in Cairene (and perhaps other dialects) of negative questions and conditionals in which *-š* is the only exponent of negation means that a natural starting point would be to assume that the stage III negative construction in Palestinian also started out restricted just to these contexts, before spreading more

widely. The data presented in sections 5.3.1 and 5.3.4, however, do not favour such a hypothesis, despite its initial plausibility. The principal finding militating against this analysis is that the restrictions on negation without *mā* in Palestinian are quite different from those operative in Cairene.

As we have seen, the stage III construction is very widely agreed to be unacceptable with perfect verbs or the pseudo-verb *ʕind-* by native speakers of Palestinian. Neither of these restrictions appears to be operative in negation with *-š* alone in interrogative and conditional clauses in Cairene (cf. exx. (150) and (153)). In order therefore to maintain the position that the situation in Cairene represents an earlier stage of the situation in Palestinian, one would have to argue that the postverbal construction in Palestinian both expanded its range to declarative clauses on the one hand, while also contracting it to exclude perfect verbs and the pseudo-verb *ʕind-* on the other. This contraction would seem hard to motivate on independent grounds.

Moreover, it is important to consider why it should be that only questions and conditionals in particular allow the omission of *ma* in Cairene negative constructions, if the progression is envisaged as being as in (161):

(161) Stage II (or stage I) negation obligatory in all clause types (earlier Cairene and earlier Palestinian) >>> stage III negation possible in just questions and conditionals (present-day Cairene and earlier Palestinian) >>> stage III negation possible in all clause types (present-day Palestinian).

Even if one were to accept, for example, the general principle of the elimination of “functional redundancy” as a mechanism driving the progression from a bipartite to a postverbal construction (as argued by Schwegler 1988: 48 for the development of postverbal negation in French), invoking such a principle to argue for the progression in (161) would seem especially problematic: there is no obvious reason why bipartite negation should be any more redundant in interrogative and conditional clauses than in others. Since a phonological explanation of (161) would be equally fraught, it is left looking like a series of changes without clear motivation. Additionally, it is unclear what role, if any, non-negative question-marking/NPI *-š* as in (151) would play in such a scenario.

If, however, we reject the progression outlined in (161) and treat stage III negation in Cairene and Palestinian as separate developments, both can be explained in terms of independently motivated changes. Moreover, an explanation of the relationship of negative and non-negative *-š* in Cairene can be given that has a direct parallel in the development of French postverbal negation, as we shall see.

5.3.5.1 NPI -š in Cairene

As we know, enclitic *-š* derives from Classical Arabic *šay?* ‘thing’. Recall from section 2.4.2 that, already in the Qur’an, there is evidence for an adverbial ‘extent’ use of the indefinite accusative *šayʔan* which is rarely attested in affirmative sentences, but is common in the scope of negation (162)-(163):

- (162) *lā yaḍurru-kum kaydu-hum šayʔan*
 NEG harm.IMPF.3MSG-you cunning-their thing.ACC
 ‘Their cunning will not harm you at all.’ (Qur’an 3: 120)
- (163) *ʔinna ʔlāha lā yaẓlimu n-nāsa šayʔan*
 indeed God.ACC NEG oppress.IMPF.3MSG the-people.ACC thing.ACC
walākinna n-nāsa anfusa-hum yaẓlimūn
 but the-people.ACC selves.ACC-their oppress.IMPF.3MPL
 ‘God does not oppress people at all; it is they who oppress themselves.’
 (Qur’an 10: 44)

It seems likely that this adverbial extent use of *šayʔan* originated through a reanalysis of the argument structure of verbs such as *ẓalama* ‘to oppress’ in (163), which typically takes a single object denoting the patient of oppression, but can also take a second object denoting the entity unjustly taken away from the patient as in *ẓalama fulānan ḥaqqahu* ‘to deprive someone of his due’. Where this kind of verb is used ditransitively, it is much more likely that the second object will be a generalizer such as *šayʔan* ‘thing’ in negative declarative, interrogative and conditional clauses than in affirmative declarative clauses. This is because it is maximally informative to deny that you have deprived someone of *anything whatsoever* (or inquire whether he has been deprived of anything, or make something conditional on his having been deprived of something/anything), whereas it is minimally informative to tell someone that you have deprived him simply of *something* without specifying what that something is (Eckardt 2006: 156). The frequent occurrence of *šayʔan* in this maximally general sense in non-assertive (i.e. negative declarative and non-

declarative) sentences and with verbs with ambiguous argument structure was likely an important contributing factor to its reanalysis as a negative polarity adverb (NPA). However, it is important to stress that in the language of the Qur'an at least (and, apparently, in Classical Arabic literature in general) this NPA use does not seem to have spread to the context of verbs with unambiguously monotransitive argument structure.⁸⁵ Thus, it could be insisted in (163) that *šayʔan* is the argument of the first token of *zalamā*, which is ditransitive (thus, 'God does not deprive the people of anything'), while the second token of the same verb is monotransitive, despite the obvious parallelism intended in the repeated use of the same verb. Similarly the verb *ḍarra* 'to harm' (or possibly 'to deprive s.o. of sthg.') in (162) and its antonym *aġnā* 'to profit', belong to the class of verbs associated with gain and loss, which crosslinguistically often feature an optional argument denoting the extent of what is gained or lost (see footnote 9). However, it is worth noting that for every single use of the verb *ḍarra* in the Qur'an where it appears to take two objects, the second of these is only ever *šayʔan*, and the verb is always negated.

Thus, already in Classical Arabic we have evidence that the ancestor of enclitic *-š* was becoming grammaticalized as an NPA. Moreover, as we also saw briefly in section 2.4.2, various dialects such as Moroccan, Libyan and (Damascene) Syrian have a clear descendant of this item in approximately this use, in what is often described as a 'question-marking' particle *šī*, for example:

⁸⁵ It also seems, in fact, to be rather infrequent in non-negative non-declarative contexts, though the following is an example:

- (i) hal zalamtu-ka min ʔajri-ka šayʔan
 Q deprive.PRF.1SG-you from pay.OBL-your thing
 'Have I deprived of your pay at all/any part of your pay?' Wensinck (1955: 224)

- (164) il-ḥawli simīn **šī**
 the-sheep fat *šī*
 'Is the sheep fat (at all)?' (Eastern Libyan; Owens 1984: 102)

In Moroccan one finds the same form (or a non-monophthongized form *šey*) in the same contexts, but given that the same form occurs in the scope of negation (165), where, as expected, it is in complementary distribution with negative *-š*, again it seems that the correct analysis is that this is an NPA as in Classical Arabic, rather than a question-marker specifically, at least in Moroccan.

- (165) ana ma nāʕəs **šey**
 I NEG sleeping *šey*
 'I'm not sleeping (at all)!' (Moroccan; Caubet 1993: 68)

The same item *šī* is also reported as a question marker for (Damascene) Syrian by Cowell (1964: 378):

- (166) ʔʔūmt-i ʔəžet mən ʕand əl-kawwa **ši**
 suits-my come.PRF.3FSG from at the-cleaner *šī*
 'Have my suits come back from the cleaner's (at all)?'

Clearly here there is no temptation to say that this use in some sense evolved from part of a bipartite negative construction, since Syrian Arabic lacks such a

construction and there is no evidence to suggest that it ever had one. Interestingly, Cowan also reports that this use of *šī* is not constrained to occur clause-finally as Owens (1984: 102), for example, reports for Libyan, but can, in fact, immediately follow the verb:

- (167) ʕam-təʔʂod šī ʔənn-i kazzāb
 PROG-intend.IMPF.2MSG *šī* COMP-I liar
 ‘Are you implying that I’m a liar?’ (Syrian; Cowell 1964: 378)

In this position, *šī* very closely resembles an alternative means of expressing questions in Libyan, namely with *-š* cliticized to the verb:

- (168) šiftū-š muḥammad
 see.PRF.2MSG-Q/NPI Muhammad
 ‘Have you seen Muhammad (at all)?’ (Eastern Libyan; Owens 1984: 102)

Despite the fact that sentential negation in Libyan (as across north Africa) is standardly expressed with the stage II construction, there is no suggestion from Owens that the *-š* enclitic in (168) expresses negation any more than *šī* in Syrian does. Instead we seem to have essentially the same situation as in Syrian, albeit that *šī* in immediately postverbal position in Libyan has become a phonologically reduced clitic.

On the basis of this comparative evidence, then, we can see that the use of *-š* (or its non-cliticized ancestor) in questions and conditionals in Cairene is part of a much

wider cross-dialectal phenomenon that is not restricted to varieties with stage II negation. Hence it seems likely that this use of *-š* is prior to, and separate from, its grammaticalization as part of the bipartite construction.

Seen in this light, the fact that negation can be expressed by means of *-š* without *ma* only in questions and conditionals in Cairene becomes a good deal clearer, especially if we consider that there has been a directly parallel phenomenon in the history of French. This has been addressed by Price (1993) and Eckardt (2006: ch.5).

Price (1993) takes issue with a statement by Ashby (1991) that is directly parallel to the apparently natural assumption discussed at the beginning of this section. Ashby (1991: 6) claims that “the dropping of the first negation [i.e. French *ne*] began in interrogative sentences”. Price (1993: 191) adds that this phenomenon is observable as early as the thirteenth century. He goes on to point out (1993: 193) that “the construction is widely attested in modern literary French in registers that are unlikely to be influenced by the popular spoken language, e.g. *Devais-je pas me servir de tout?* (de Gaulle, *L’Appel*, 1954: 270) which is only susceptible of a ‘negative’ interpretation (‘Was it not my duty to use all possible means?’)”. It does not follow, however, that such uses of *pas* (and, earlier, the now obsolete *point* and *mie*) without *ne* are directly descended from a negative construction with *ne*, or that the widespread phenomenon of *ne*-omission in present-day colloquial French necessarily had its origins in *pas/point/mie* without *ne* in negative interrogatives and conditionals many centuries earlier. Rather, what we find is that *point* and *mie* and possibly also *pas* (but see Eckardt 2006: ch.5) grammaticalized first as NPAs, and are thus commonly found in affirmative interrogative and conditional clauses (in addition to

negative declarative clauses), before they came to be understood as inherently negative. Price's (1993: 194) claim is then that:

“as negative particles [e.g. *pas*, *point*, *mie* etc. – CL] came to be increasingly used in negative constructions, at the expense of *ne* alone, it is not surprising that they came to take on a negative value themselves, first of all in negative clauses (as a result of which, in course of time, the negative value was transferred from *ne* to the second particle as in *je sais pas*), and that this was carried over into the interrogative construction we have been discussing, which originally had no negative value at all.”

This is precisely the development I propose to reconstruct for *-š* without *ma* in questions and conditionals in Cairene.⁸⁶ We have seen evidence for the development of *šay?* into an NPA, which is definitely not itself negative, already in Classical Arabic, as well as in a range of present-day dialects, where it may in some cases have become restricted in its distribution just to questions. My claim is then that at least in Cairene the subsequent grammaticalization of *šay?/ši* as the second element of the bipartite negative construction led to the partial reanalysis of NPA *-š/ši/šay?* such that it could at least optionally express negation. In interrogatives this optionality remains, but it appears that *-š* without *ma* in conditionals is now always interpreted as negative. Although it is not entirely clear why the negativization of NPAs should proceed in this order (conditionals before interrogatives), note that this also has a direct parallel in French. In the present-day standard language *jamais* can be used to

⁸⁶ See also Willis (forthcoming: section 5) for a directly parallel proposal concerning the development of postverbal negation in Middle Welsh.

mean ‘ever’ in affirmative interrogative clauses, but it cannot be used in affirmative conditional clauses (only in negative ones together with *ne* to mean ‘never’) outside of the frozen expression *si jamais* ‘if by any chance’ (Hansen 2009: 14-15). The proposed reconstruction of the development of purely postverbal negation in Cairene is summarized in (169):

(169) *šay?* grammaticalizes as an adverb ‘at all’ restricted to negative, interrogative and conditional contexts (pre-diaspora spoken Arabic) >>> *šay?/ši/-š* in negative contexts is reanalysed as part of a bipartite negative construction (early north African Arabic) >>> *-š* becomes inherently negative, at least potentially, wherever it occurs, thus capable of expressing negation on its own in questions and conditionals, though *ma* remains obligatory in negative declarative clauses (early modern Cairene) >>> *-š* obligatorily expresses negation in conditionals, optionality between negative and affirmative interpretation of *-š* without *ma* remains in interrogatives (present day Cairene).

Despite the apparent lack of theoretical parsimony involved in positing two separate developments, one in Cairene and one in Palestinian, leading to what looks superficially like the same postverbal negative construction, rather than one as illustrated in (161), I have tried to show that there are two compelling reasons to favour the scenario in (169). First, it avoids the necessity of accounting in an *ad hoc* fashion for how the contextual restrictions on the postverbal construction came to be so different in the two dialects, and second it provides a clear explanation for why negation by means of *-š* without *ma* in Cairene is possible only in interrogatives and

conditionals. Accounting for the rather different restrictions on stage III negation in Palestinian forms the topic of the next subsection.

5.3.5.2 Stage III negation in Palestinian

As we saw above, sentential negation in Palestinian can be expressed by means of the stage III construction with the imperfect (both with and without the *b-* prefix) of regular verbs as well as with the pseudo-verbs *fī*, *maʕ-* and *bidd-*, but not with the perfect of regular verbs, nor with the pseudo-verb *ʕind-*. We also saw that this distribution cannot be fully accounted for by means of a synchronic phonological rule whereby an underlying *mā* is deleted always and only before a following labial consonant.

Nevertheless, the correlation between the presence of a following labial and the contexts in which negation without *mā* is grammatical remains compelling, despite not being total. My proposal, then, in keeping with the above-mentioned assumptions of internal reconstruction, is that the restrictions on the stage III construction today can give an indication of the context in which postverbal negation first originated in Palestinian, and how it spread from that context to achieve its present distribution.

Given that the stage III construction is not available for the whole paradigm of regular verbs (specifically not the perfect tense) it seems unlikely that this is the context in which the construction originated. Instead, let us examine the possibility that it originated among the pseudo-verbs.

Recall that a statement of the distribution of the stage III construction in terms of *mā*-deletion before a labial does make the correct predictions for grammaticality if we restrict its application just to the pseudo-verbs. Note also that this type of phonetic

reduction is most likely to happen with the most frequently occurring lexical items, which have been abundantly demonstrated to favour rapid and reduced articulation (cf. Schuchardt 1885, Bybee and Hopper 2001: 10-13, Bybee 2003 and many in between). There is no doubt that the pseudo-verbs are very frequent relative to regular verbs, as are their translation equivalents in other languages 'have', 'want', 'there is'. Thus, even if the reduction of *mā* to zero is limited in its occurrence to just the highest-frequency verbs, three out of four of the commonest pseudo-verbs in Palestinian will be affected. The fourth pseudo verb *ʕind-*, however, will not be affected despite its similarly high frequency, since it begins with a pharyngeal rather than a labial. Nevertheless, if a situation is established whereby speakers are perceived as routinely omitting *mā* with at least these three items, then a potential problem for successful acquisition of the old system of negation arises.

For individuals acquiring Palestinian at this stage, the conservative analysis of input of this kind would be to assume that sentential negation is always underlyingly expressed by means of the stage II construction,⁸⁷ and that the first element of this construction is then occasionally subject to deletion as a result of the gradual reduction to zero of the magnitude of the articulatory gestures associated with *mā*, just in the context of these highest frequency labial-initial items.

On the other hand an alternative, more innovative analysis of this input would be to posit a fully-fledged stage III construction, not derived by phonological deletion of *mā*, that exists as an option alongside the bipartite construction, albeit restricted in its occurrence to the pseudo-verbs (as well as possibly the *b*-imperfect of the highest frequency regular verbs). The impossibility of *ʕind-* with this new construction would

⁸⁷ I leave the stage I construction out of consideration for present purposes.

then have to be learnt as a lexical exception – something that would presumably not cause much of a problem, given, again, the frequency of *ʕind*-⁸⁸ and the availability for this verb of the alternative periphrastic negative construction (*mā*) *fi-š ʕind*-.

My proposal, then, is that the development of the stage III negative construction in Palestinian was in fact triggered by the phonetic reduction to zero of *mā*, but only in the context of the highest frequency labial-initial verbs, of which some of the most frequent of all would have been the pseudo-verbs (except *ʕind*-).

It is perhaps not too far-fetched to suggest that the pseudo-verbs should have had a negative construction that was essentially unique to them, given that they have a range of other irregular properties, which strongly mark them out as a set distinct from the regular verbs. Nevertheless, a less restrictive analysis of the input generated by deletion of *mā* would perhaps be more natural. We can envisage this as follows.

First of all, note that any lexical item can be viewed as a member of several more or less restrictive grammatical sets. Hence the pseudo-verbs may be seen as constituting a highly restricted set unto themselves, or as part of the much larger set of all verbs. However, one of the defining properties of the pseudo-verbs is, as noted above, that they lack a non-periphrastic perfect tense. Thus they can also be viewed as members of the intermediate set of all non-past verb forms. It seems natural to suppose, then, that individuals acquiring Palestinian Arabic in the relevant period could just as well make the generalization from the input they were exposed to that the stage III construction was available not just for the pseudo-verbs, but for this intermediate set of all non-past verb forms. Such a generalization would result in the

⁸⁸ Frequency and irregularity have also been shown to be strongly correlated (Bybee 2003).

distribution of the postverbal construction that we observe today: ungrammaticality with the perfect tense (including of verbs beginning with labials), but grammaticality with the imperfect, both with and without the indicative *b*-prefix and including prohibitives, and also with the pseudo-verbs. *šind-* would then continue to have to be learnt as a lexical exception to this generalization.

Before we go further, a word is in order here on issues of learnability and (over)generalization in acquisition. It has long been recognized that, given that children do not have access to negative evidence in acquiring the grammar of their first language, the explanation of how they nevertheless converge on the target grammar in such a high proportion of cases is made considerably easier by assuming that acquisition is guided by the Subset Principle (Berwick 1985: 235-238). In essence, this principle states that “the learner must guess the smallest possible language compatible with the input at each stage of the learning procedure” (Clark and Roberts 1993: 304-5). Following this strategy entails not making overgeneralizations in the course of acquisition whose incorrectness could not be demonstrated by means of positive evidence (and innate principles of grammar) alone. Clearly this is an eminently sensible way to proceed. However, there must be some doubt as to whether the acquisition path proposed in the previous paragraph is in conformity with the Subset Principle. If it is not, then perhaps it is implausible? In fact I want to argue that, despite initial appearances, my proposal does not contravene the Subset Principle on a reasonable interpretation of what the principle means in practice.

This being said, it is by no means certain that acquirers do in fact always stick rigidly to the Subset Principle – see Bowerman (1988) for a wide range of attested

overgeneralizations that children make and must, therefore, have some means of later retreating from. Moreover, as Fodor and Sakas (2005: 514) point out, “it is salutary to remind oneself every now and then, when invoking [the Subset Principle], that there is as yet no satisfactory theory of what work it should do, or how, even in principle, it should do it.”

The difficulties arise when one considers that generalization in acquisition is vital: an overcautious adherence to the Subset Principle – for example, never assuming that the syntactic properties observed for n lexical items generalize to a whole class – will result in the failure ever to move beyond a tiny subset of the target grammar (Fodor and Sakas 2005: 516, Albright 2008: 151).

What mechanism(s) children have for retreating from attested overgeneralizations (such as *I said her no* [= told her], *Don't giggle me* [= make me giggle], *Why is the laundry place stayed open all night* [= kept], etc. (Bowerman 1988: 79)) is unclear (see Randall 1992 for one suggestion), but it seems likely that these mechanisms are not as secure as sticking to the Subset Principle in the first place. That is, one would expect them to fail at some point and for syntactic change to result. In fact, it would appear that this kind of scenario has been uncontroversially assumed in the literature on syntactic change since Timberlake's (1977) influential article on reanalysis and actualization, such that “the gradual mapping out of the consequences” of a reanalysis (Timberlake 1977: 141) is essentially equivalent to successive failures to retreat from overgeneralizations with respect to the input. We take up this point again in section 5.3.6.

Returning now to the acquisition path proposed above for earlier Palestinian negation, what I am arguing for is a reanalysis followed by an extension, in the

senses of Harris and Campbell (1995), where reanalysis is defined as “a mechanism which changes the underlying structure of a syntactic pattern and which does not involve any modification of its surface structure” (1995: 50) and extension as “a mechanism which results in changes in the surface manifestation of a pattern and which does not involve immediate or intrinsic modification of underlying structure” (1995: 51). The reanalysis consists of acquirers taking a string such as (170) and, instead of analysing it as underlyingly involving a stage II construction whose preverbal element is unpronounced for phonological reasons, they take it at face value and posit a stage III construction underlyingly. Thus we have a change in underlying structure with no modification at the surface.

(170) *biddī-š* *maṣāri*.
 want.1SG-NEG money
 ‘I don’t want money.’

For the reasons given above, we cannot now expect overly strict application of the Subset Principle. Specifically, it is unlikely that an acquirer would initially posit a negative construction that is restricted just to the lexical items with which it is found in the input, namely the three pseudo-verbs *fi*, *maṣ-* and *bidd-* and perhaps a handful of the most frequent regular verbs in the *b*-imperfect. Rather, a plausible, highly conservative generalization to make on the basis of this input would be, as suggested above, to assume that the stage III negative construction is an option just for non-past verbal sentences: an extension to “a natural class based on categories already relevant to the sphere in which the rule applied before it was extended”

(Harris and Campbell 1995: 101). Many much less restrictive (over)generalizations could be envisaged, but this seems to be the most natural one consistent with the spirit of the Subset Principle. Of course, initially most acquirers would have later successfully retreated from this generalization in the same way that children acquiring English retreat from the overgeneralizations such as applying the dative alternation to verbs such as *say*. Given the lack of security of generalizing and retreating versus sticking rigidly to the Subset Principle, however, one would expect individual failures to retreat to have built up over time, until the current distribution of stage III negation became a widespread and accepted feature of the grammar of Palestinian. In fact, we will examine some evidence in the following section which suggests that children acquiring present-day Palestinian have begun making still less restrictive generalizations, albeit they appear to be successfully retreating from them for the time being.

The advantage of the present analysis, summarized in (171), is that it offers an explanation of the development of stage III negation in Palestinian which is based on independently motivated mechanisms of syntactic change, and which takes synchronic restrictions on the distribution of the construction not as exceptions to be glossed over, but as important clues to how this construction originally came into being.

(171) Stage III negation not possible, only stage II (pre-Palestinian 1) >>> stage III negation appears on the surface as a result of phonological deletion of underlyingly present *mā* before very frequent labial-initial pseudo-verbs (pre-Palestinian 2) >>> appearance of stage III negation in this context is

reanalysed as a syntactic reality (pre-Palestinian 3) >>> stage III negation is extended (generalized) to all non-past verb forms, except *ʕind-* (present-day Palestinian) [>>> stage III negation is extended to all verbs in all tenses (possible future Palestinian)].

5.3.6 Future prospects

Having seen the current distribution of the stage III construction in Palestinian in the light of extension from the pseudo-verbs to the set of all non-past verb forms, it is tempting to speculate about the prospects for further extension to a more general set: that of all verbs in all tenses.

As mentioned in section 5.3.4, the single informant who judged the postverbal construction with a perfect verb to be acceptable volunteered the information that her response was influenced by the types of structures she recalled her young children producing. In fact, approximately one third of my informants volunteered the same information: that strings such as (156) are not acceptable as far as they were concerned, but are reminiscent of the kinds of “mistakes” that one often hears children making.

Some empirical support for these anecdotal observations comes from a dissertation on the acquisition of negation in a variety of Jordanian Arabic that closely resembles Palestinian with respect to negation (Abu El-Haija 1981). Unfortunately the author does not specify precisely which dialect he is studying the acquisition of, nor does he mention any tense split in the distribution of the stage III construction in the target grammar of this dialect. He does state, however, that “from the results for split negation [*mā... -š*]... it is one of the forms acquired late in [children’s] language

development process. Younger children have difficulty employing this form. We can say that it is almost absent in their spontaneous speech.” (Abu El-Haija 1981: 106). The relevant results are shown in Table 5.16 (adapted from Abu El-Haija 1981: table 7), which clearly shows the stage III construction dominating for children under four years old.

Table 5.16 Tokens of negative constructions in the speech of young acquirers of Jordanian

Forms	Age 2.5 years			Age 3.5 years		
	Child 1	Child 2	Child 3	Child 4	Child 5	Child 6
<i>mā</i> only	0	1	0	1	1	1
<i>mā...-š</i>	0	0	1	0	0	1
<i>-š</i> only	8	6	9	0	6	4

Unfortunately no information is provided as to whether some of these tokens of stage III constructions are overgeneralizations to the perfect tense. However, the near absence of tokens of stage II constructions would suggest that this is a possibility, particularly if we compare the figures from the mothers of these children, shown in Table 5.17 (adapted from Abu El-Haija 1981: table 8):

Table 5.17 Tokens of negative constructions in the speech of Jordanian mothers

Forms	Mother 1	Mother 2	Mother 3	Mother 4	Mother 5	Mother 6
<i>mā</i> only	2	6	1	4	4	4
<i>mā...-š</i>	4	7	5	7	12	5
<i>-š</i> only	2	3	2	1	5	7

Here we see that for every mother of the children shown in Table 5.16 except Mother 6, tokens of the stage II construction outnumber tokens of the stage III construction, thus showing the opposite pattern to their children. Hence the speech of the children appears to be target-deviant at least in the relative frequency of use the stage II and stage III constructions.

Although this rather circumstantial evidence is far from being conclusive, the picture that emerges is that apparently children acquiring the mixed negation system of the dialects spoken in and around Palestine today initially seem to make a rather broad generalization concerning the distribution of the stage III construction. That is, they view the lexical items in their input that are negated with the postverbal construction as members of the set of all verbs (in all tenses). This generalization must then later be scaled back so that the stage III construction is restricted to the set of all non-past verb forms (and excluding the pseudo-verb *šind-*). It seems natural to suppose, then, that successive generations of children acquiring this system will gradually fail to learn and apply this restriction, and the stage III construction will thus be extended throughout the whole of the verbal paradigm. More research specifically on the acquisition of negation in Palestinian and neighbouring Arabic varieties would

be necessary to verify whether the initial overgeneralization of the stage III construction is as widespread as I have speculated here.

To conclude, section 5.3 has proposed a reconstruction of the origin of the postverbal negative construction (-š without preverbal *mā*) as it appears in both Palestinian and Cairene Arabic. I have claimed, on the basis of empirical data collected from grammars and fieldwork, as well as comparative data from the history of French, that postverbal negation developed separately in these two dialects. In Cairene my proposal is that it developed out of a non-negative use of -š in questions and conditionals (NPI contexts). In negative declarative clauses *ma* was always present and continues to be obligatory today. In Palestinian my proposal is that stage III negation was triggered by phonological processes affecting the most frequent verbs, and was then generalized to the set of all non-past verb forms. Whether it will be further extended to the whole verbal paradigm remains to be seen.

6 Developments in the negation systems of other Afro-Asiatic languages

This chapter explores developments in the negation systems of a number of Afro-Asiatic languages not discussed in detail thus far, and analyses them, insofar as the data permit, using the principles and methodology established in previous chapters. We will see that while many of the developments are superficially familiar, including several instances of the development of bipartite negation, the input structures for the changes in question are often rather unexpected, given what we have seen up to this point. Moreover, when we turn to consider the development of negative concord, we will see that the presence of this phenomenon in Ge'ez, when considered alongside its absence in the vernacular Ethiopian Semitic languages and all stages of Akkadian and Egyptian-Coptic, leads us to question some widely held assumptions about what pressures can promote the development of negative concord and what, in turn, can impede it.

The chapter is structured as follows. After a brief discussion in section 6.1 of the small amount of previous literature on the subject of negation in Afro-Asiatic languages generally, section 6.2 explores a number of Jespersen-type developments in the Semitic and Cushitic languages of the Horn of Africa. Section 6.3 then deals with the development of negative concord in Afro-Asiatic languages other than Arabic, with a primary focus on Hebrew.

Since a number of the languages discussed in this chapter are not well-known to non-specialists, I provide here some simplified trees for the relevant genera and sub-groups (cf. Faber 1997; Lamberti 1991).

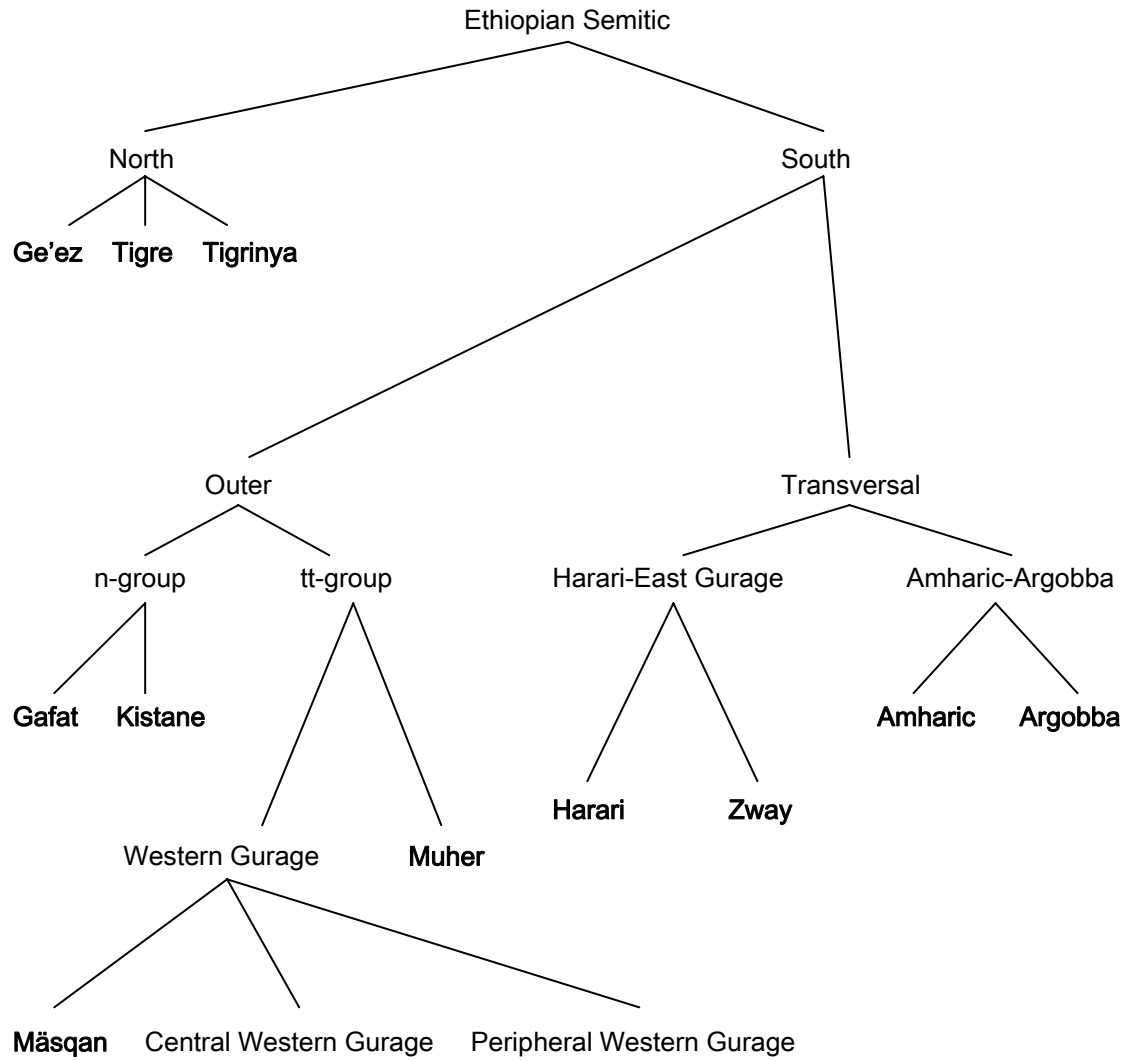


Figure 6.1

Ethiopian Semitic family tree

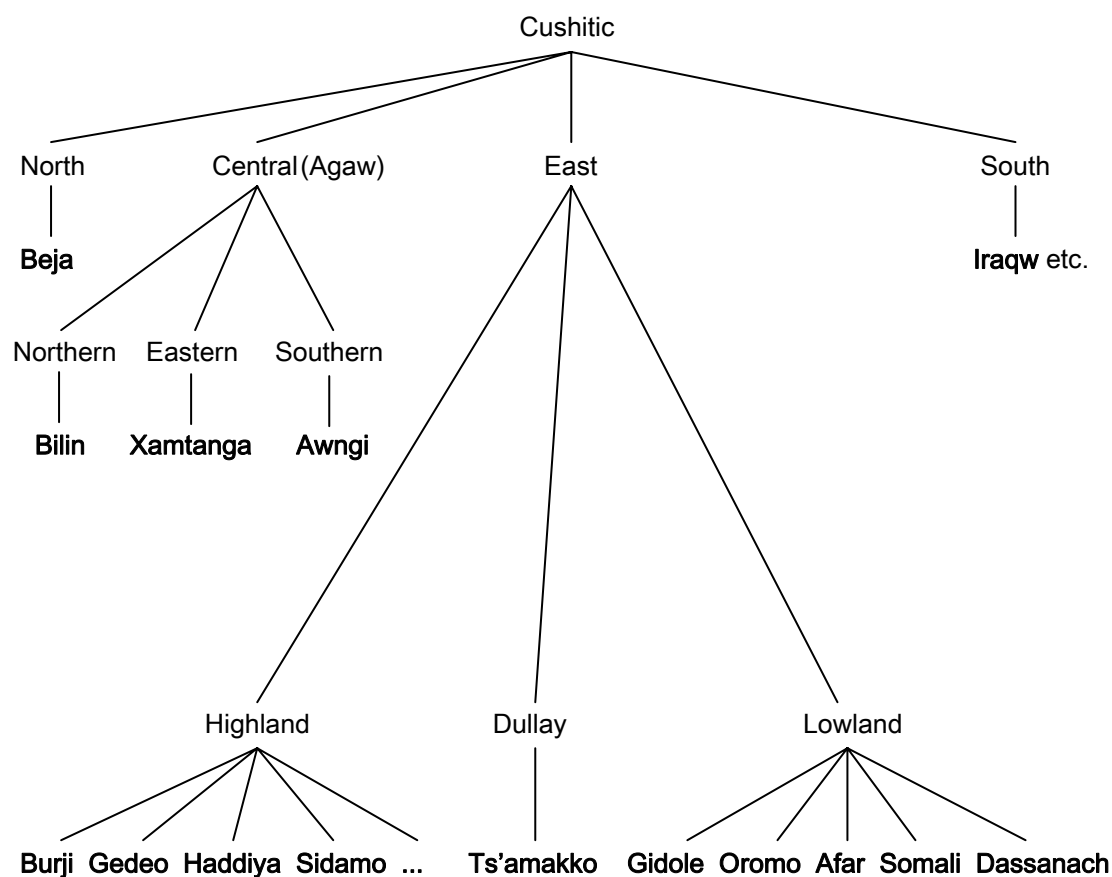


Figure 6.2

Cushitic family tree

6.1 Previous literature

The only previous literature I am aware of that has as its central concern the historical development of negation in a range of Afro-Asiatic languages is by Faber (1988, 1991). Although her focus is not on Jespersen-type developments, it is worth considering her claims in the light of some of the topics we have looked at in the rest of this thesis.

Faber (1991) attempts to reconstruct the negative and interrogative markers of Proto-Afro-Asiatic using a methodology that has much in common with the

'megalocomparison' school (e.g., Greenberg 1987, Ruhlen 1994). She states (1991: 41) that her article "has its origin in a trivial observation: the interrogative pronouns and adverbs of the Semitic languages... bear a striking resemblance to one of the negative markers. The interrogative words begin in *m-* and the NEG is *mā*." While interrogatives in *m-* are indeed common in Semitic and wider Afro-Asiatic, it is far from clear that the statement concerning negation is valid as it stands. Within Semitic it is only in Arabic that there is uncontroversially a negator (as well as an interrogative pronoun) *mā*.

Faber's claim that a similar negator *ma* is found in relic expressions in Biblical Hebrew is problematic. The standard grammars and dictionaries do not recognize a negative function for this item. It seems likely that this claim is based on an idiosyncratic analysis of certain obscure passages containing the Hebrew wh-pronoun *ma* (cf. Lipínski 1997: §47.15, Rubin 2005: 50).

Faber also points to the enclitic negator *-m* in certain Ethiopian Semitic languages. We will see in section 6.2, however, that this represents a (relatively speaking) recent innovation within Ethiopian Semitic that appears to have its origin in the reanalysis of an enclitic conjunction. It is unlikely to be derived from a Common Semitic negator.

Within Semitic, the only other evidence that Faber adduces in support of her "trivial observation" is that Harari (Ethiopian Semitic) has a word *mē?* 'no' and a negative copula *elum* (this is in fact the negation of the existential verb *hāl* featuring the same enclitic negator *-m* as mentioned above; Wagner 1997: 507), and that Akkadian (East Semitic), Hebrew, Phoenician and Xanthos Aramaic (Northwest Semitic) all have NPIs featuring one or more bilabial nasals: *mimma* 'anything',

məʔuma ‘anything’, *mnm* ‘anything’ and *mtwm* ‘ever’, respectively. The oddity of listing these NPI indefinites under the heading of negative elements is partially explained in a (1988) paper, in which Faber shows that these items are indeed NPIs, and that they all probably involve an original generalizing *-ma* suffix which can be reconstructed to Proto-Semitic, and which, based on no more evidence than that just discussed, Faber derives from a hypothesized Proto-Semitic negator **ma*.

As Haspelmath (1997: 231) points out in his brief discussion of this article, this scenario runs counter to a sensible “general rule of diachronic typology that reconstructed changes can never disprove a proposed universal”: overwhelmingly, empirically attested developments in indefinite pronoun systems are from less to more negative, and not vice versa. Given that the only certain case of negative *mā* is in Arabic, it is far from clear that one can reasonably talk of a common Semitic negator *mā*, and thus especially unlikely that this should be the source of the generalizing suffix *-ma*.

Despite these problems within Semitic, Faber (1991) also points to the presence of a prohibitive *m-* proclitic in Egyptian and a negative suffix *-ma* in Hamar (which is apparently the only Omotic language to have this item; see footnote 1 on the doubt as to whether Omotic languages should even be classified as part of Afro-Asiatic), and lists a number of Lowland East Cushitic languages which all have a proclitic *ma-* negator (presumably present in Common Lowland East Cushitic). On this basis she states, without obvious justification, that there is “ample evidence that the Sem[itic] NEG *mā* was inherited from A[fro]-A[siatic]” (Faber 1991: 412).

Given that wh-pronouns with initial *m-* are widespread in Semitic, Berber, Chadic and Cushitic, and are also found in Egyptian, it does seem plausible to reconstruct a

wh-formative in *m-* to Proto-Afro-Asiatic as Faber (1991: 412) suggests. In view of the above, however, Faber's main proposal to explain the resemblance she perceives between Afro-Asiatic negative and interrogative markers is hardly warranted. She suggests, albeit "tentatively" (1991: 420) that "in all of A[fro]-A[siatric] but Om[otic], a word meaning 'what?' developed from prefixation of a [question particle] derived from the NEG **(m)ba* to a word meaning 'thing'". She does not explain how or why a question particle should have developed from a negator, nor what the status of the parentheses in her reconstructed negator **(m)ba* is, nor what the form of the word meaning 'thing' might have been. Given all this, it seems unwise to treat her reconstruction as definitive. Rather, it seems more likely, a) that the Semitic generalizing suffix is straightforwardly derived from an *m-* interrogative (a change also attested by, e.g., Russian *koe-*, Haspelmath 1997: 232), and b) that the two certain instances of homophony between interrogative and negative markers in *m-* – in Arabic and Lowland East Cushitic – represent separate instances of the change wh-pronoun > negator, for which I sketched a possible scenario in section 1.3.1.⁸⁹ Hence the resemblance of the Arabic and Cushitic negators is unlikely to be due to common inheritance from Proto-Afro-Asiatic, while the resemblance of their interrogative pronouns could well be. (See Pat-El forthcoming for further objections to Faber 1991).

Another relevant piece of research is the recent article by Van Gelderen (2008), mentioned in section 2.2. This article is concerned with various negative cycles crosslinguistically. It features a section on Afro-Asiatic languages with bipartite

⁸⁹ Recall from section 2.3.1 that this change is also separately attested by Spanish Arabic *iš* (< *aš* 'what' < *ʔayy šayʔ* 'which thing').

negative constructions, giving data on Central Atlas Tamazight, Kabyle, Tarifit (Berber), Standard and Moroccan Arabic, Zway, Amharic (Semitic), Koorete (Omotic), Somali, Beja (Cushitic), and Hausa (Chadic).

Among these data Van Gelderen identifies three main diachronic developments. The first is Jespersen's Cycle, exemplified (among the languages in Van Gelderen's sample) by the above-mentioned Berber languages as well as Moroccan Arabic.

The second development Van Gelderen mentions is from an interrogative marker to a preverbal negator occupying the head of PolP: the last-mentioned change in the discussion of Faber (1991) above. She notes this development for Arabic (following Rubin 2005: 50), and also for Somali (Lowland East Cushitic), though, as I pointed out above, this development probably took place already at the stage of Common Lowland East Cushitic. In addition to the preverbal negator *má*, Van Gelderen notes that Somali negative sentences also involve a special negative form of the verb. She hypothesizes that this special negative form is the remnant of an earlier negator that predates the introduction of *má*. We will see in section 6.2 that this is unlikely to be correct: rather, *ma* (with or without high tone) is the original negator in Lowland East Cushitic and the special negative form of the verb is derived from a former perfect tense negative auxiliary that is still identifiable as such in the related language Afar. This is in fact, then, a case of the third development that Van Gelderen identifies: that of a negative auxiliary from a lexical verb, as in Koorete (Omotic) *ba* < 'to disappear' and Beja (North Cushitic) *rib* < 'to refuse'. This particular development is relatively common crosslinguistically (Van Gelderen identifies it as a central feature of the history of negation in a number of Uralic languages for example), but to the extent

that it occurs in Afro-Asiatic languages beyond Lowland East Cushitic there is insufficient diachronic data to comment further here.

Van Gelderen also gives data on negation in a couple of Afro-Asiatic languages without making suggestions as to how the structures in question developed. Concerning Amharic and Zway (a related Ethiopian Semitic language not usually thought of as a variety of Amharic as Van Gelderen labels it; see Figure 6.1), she notes that these languages feature a bipartite negative construction that appears structurally similar to that of Arabic: there is a proclitic on the verb *a/-* (in both languages) and an enclitic *-u* in Zway, *-m* in Amharic. We will see in greater detail in the following section that this *-m* enclitic (as well as its vocalized Zway cognate, and cognates in other Ethiopian Semitic languages) has its origin in the contact-induced reanalysis of an ‘and’-conjunction.

Additionally, Van Gelderen gives data on negation in Hausa (Chadic), which again features a bipartite construction as shown in (172) (= Van Gelderen’s (76)), though here the second element occupies a clause-final or clause-late position and, except for the tonal contrast, is identical to the preverbal negator.

(172) **bà** kà kāwō àbinci **ba**

NEG you bring food NEG

‘You didn’t bring food.’ (Hausa; Kraft and Kirk-Greene 1973: 38)

Van Gelderen makes no suggestion as to the origin of this construction, but these data look similar to the ‘resumptive’ negative construction found in Afrikaans (den Besten 1986; Bernini and Ramat 1996: 51-81; Biberauer 2008), Brazilian Portuguese

(Schwenter 2006), certain northern Italian dialects (Parry 2009) and, as we saw in section 2.4.3, arguably Modern South Arabian. As such, we might speculate that here too the path of development for the Hausa construction was similar to that suggested for Brazilian Portuguese by Schwegler (1988), whereby a second negator routinely appended to a negative sentence for emphasis (as in *I don't like that, not (at all)*) is reanalysed as functioning as part of the negation of the main clause. In the absence of diachronic evidence, however, this must remain a speculation.

6.2 Jespersen-type developments in the Horn of Africa

As noted above, a bipartite negative construction is found in several Ethiopian Semitic and Lowland East Cushitic languages. These are addressed in sections 6.2.1 and 6.2.2 respectively.

6.2.1 Bipartite negation in Ethiopian Semitic

Many (but not all) of the Ethiopian Semitic languages have a bipartite negative construction whose precise form varies slightly from language to language, but whose syntax is essentially identical in all of them. Thus in Tigrinya (North Ethiopian Semitic) we have a proclitic *?ay* and an enclitic *-n*:

- (173) nəḥna mənəm ḥadä nägär **?ay-gäbärna-n**
 we any one thing NEG-do.PRF.1PL-NEG
 'We didn't do anything.' (Tigrinya; Kogan 1997: 442)

In Ge'ez and Tigre, the other two North Ethiopian Semitic languages, we have just a stage I construction with *?i* (a reduced form of *?ay*).⁹⁰ In Amharic and Harari (Transversal South Ethiopian Semitic) negation is expressed with a proclitic *a/-* and an enclitic *-m*:

(174) **al-näggärku-m**

NEG-tell.PRF.1SG-NEG

'We didn't tell.'

(Amharic; Hudson 1997: 471)

In Zway (Outer South Ethiopian Semitic) we have proclitic *a/-* (often reduced to *a-*) and enclitic *-u* (reduced to zero in forms with an original final vowel):

(175) **a-yfärək-u**

NEG-be.patient.IMPF.3MSG-NEG

'He is not patient.'

(Zway; Leslau 1999: 75)

That the Zway enclitic *-u* is the result of vocalization of original /m/ is suggested by the intermediate position of Argobba (Transversal South Ethiopic), in which the enclitic negative has the original form /m/ following /u/, but /w/ following other vowels, /aw/ following nasal consonants and /u/ following other consonants (Hudson 1997: 471-2).

Thus the Ethiopian Semitic languages can be divided into three types with respect to the expression of negation: those that have stage I negation only, those

⁹⁰ An exception to this is the Rigbat dialect of Tigre, which, according to Elias (2007), has a bipartite negative construction *y(V)...-n(ni)* that appears to be a borrowing from Tigrinya.

that have a stage II construction with an enclitic *-m* or a derivative thereof, and Tigrinya, which has stage II negation with an enclitic *-n*.⁹¹ Representatives of the first type (i.e., languages which have been conservative with regard to the expression of negation) are found in all branches of Ethiopian Semitic, suggesting that the stage II construction cannot be explained by means of a single innovation in a common ancestor.

Historical records of Ethiopian Semitic languages can shed a limited amount of light on the development of the stage II construction(s). The Ethiopian Semitic language with by far the oldest and most extensive written attestation is Ge'ez, but this has been conservative with respect to Jespersen's Cycle. The only other languages for which we have records going back more than about a century are Amharic and Harari. The oldest extensive and easily accessible record of Amharic is Ludolf's (1698) grammar. Here we find precisely the same situation as today: a bipartite construction *a/-...-m* found with all main clause verbs (Ludolf 1698: 52).

Harari is attested in documents from approximately the mid-eighteenth century or possibly earlier (Wagner 1997: 486). Interestingly, the situation in the oldest Harari documents is quite different to what obtains in Modern Harari: in the eighteenth-century texts collected in Wagner (1983), out of the 69 instances of negation that I counted in contexts where a stage II construction with *-m* would be obligatory in Modern Harari, all but four lacked a negative *-m*. An example of each variant is given in (176).

⁹¹ An exception to this generalization is the Peripheral Western Gurage dialect cluster (Outer South Ethiopic) which features a bipartite construction *a-...-ka/-ta/-da* (Hetzron 1997b: 545). It is unclear what the origin of this second element is.

- (176) (a) **al-xätäre-w** **zi-jañ** **bär?i**
 NEG-prevent.PRF.3MSG-him of-exalted gate
 ‘The gate of the Exalted One did not block his way.’
 (Old Harari; Wagner 1983: 59)

- (b) **al-qäbäṭa-m** **mauṣidata**
 NEG-fail.PRF.3MSG-NEG promise
 ‘He did not break his promise.’
 (Old Harari; Wagner 1983: 79)

Additionally, there are a handful of examples, such as (177), where there is an *-m* enclitic present, but it is difficult to say whether this is the negative *-m* or a different, homophonous item that functions as a conjunction. In considering (177) note that asyndetic coordination is common in Ethiopian Semitic languages, also that in Modern Harari the position of negative *-m* is variable and need not cliticize to the verb, and that the negative of the existential verb *ḥal* is an irregular form *ēl(um)* (Wagner 1997: 502).

- (177) **ḥoji** **bi-dinät** **ge-m** **ēl-bä-na** **way, geš**
 today in-property world-and/NEG there.is.NEG-in-us woe tomorrow
bi-āxirat-um **ēl-bä-na** **way**
 in-end-and/NEG there.is.NEG-in-us woe
 ‘We do not suffer (have) misery on Earth today, neither will we suffer (have) misery in the hereafter tomorrow.’
 (Old Harari; Wagner 1983: 214)

Examples such as (177) raise the possibility of a diachronic link between the enclitic *-m* negator in Modern Harari and the homophonous conjunction, and this impression is strengthened by the fact that Amharic, in addition to its identical enclitic *-m* negator, also has an identical enclitic *-m* conjunction (Leslau 1995: 882).

I am not aware of any other language in which a new negator has developed from a former conjunction, which ought to make us wary of adopting this hypothesis in the present case. Indeed, at first glance, the Tigrinya negative enclitic *-n* appears to be problematic, given that it cannot be cognate with *-m* in other Ethiopian Semitic languages (Proto-Semitic **m* and **n* are preserved unchanged in all Ethiopian Semitic languages). However, in Tigrinya we find that the commonest 'and'-conjunction is not the same as in Harari and Amharic, but is in fact a different item that is homophonous with the Tigrinya negative enclitic *-n* (Kogan 1997: 442).

This starts to make the conjunctive-origin hypothesis a good deal more plausible, and we find a link between negative and conjunctive clitics in Ethiopian Semitic suggested already in Hetzron (1972: 94-98). However, Hetzron went further and hypothesized that this development was not internal to Semitic, but was in fact the result of contact with Agaw languages (Central Cushitic). Hetzron (1972: 98) points out that in Awngi (a Southern Central Cushitic language spoken by several hundred thousand people southwest of Lake Tana, and undoubtedly an important contact language for Amharic and Tigrinya in particular) there is a bipartite negative construction in which the second element is an enclitic *-ki*. The crucial point is that Awngi also features an identical enclitic conjunction. Thus we appear to have here another case of polysemy (or, perhaps better, homophony) copying (Heine and

Kuteva 2003, 2005; cf. section 3.4.2.2): the employment of a sound string that serves both as an 'and'-conjunction and as part of a negative construction has been transferred from Awngi to a number of different contact languages, each of which has reanalysed the form ordinarily used for coordination such that it can also function as a negator. Importantly, negative *-kí* in Awngi is possible only in matrix clauses, and precisely the same is true of its counterpart *-m* in Amharic and Harari (Hudson 1997: 471, Wagner 1997: 503).

Too little is known about the sociolinguistic situation in Ethiopia in the period when this transfer would have taken place to speculate about whether the agents of change in this case would have been native speakers of Awngi or of the relevant Ethiopian Semitic languages, or perhaps a combination of the two. However, a clear advantage of Hetzron's contact explanation for Jespersen's Cycle in Ethiopian Semitic is that it can explain its sporadic occurrence in all branches of the Ethiopian Semitic family tree, as well as the formal difference of the bipartite construction in Tigrinya versus the other languages discussed above.

We have a similar situation here as we saw for Jespersen's Cycle in Arabic in chapter 2: an internal explanation based on reanalysis in contexts such as that given in (177) is fine in principle (though, as we observed above, the development conjunction > negator is hardly a commonly observed grammaticalization path). However, here, as in Arabic, an internal explanation leaves us unable to explain the uneven distribution of the relevant construction across a range of closely related languages. By contrast, the contact-based explanation makes a clear and potentially falsifiable prediction: Jespersen's Cycle will only have occurred in those Semitic languages which were in intensive contact with Southern Central Cushitic languages,

as well as those languages to which it could have diffused more recently through contact with Amharic. Bipartite negation in Harari, a language which has historically been surrounded by Lowland East Cushitic, rather than Central Cushitic languages, and which seems to have developed the bipartite construction only relatively recently, would appear to be a case of the latter.

6.2.2 Bipartite negation in Lowland East Cushitic

A number of Lowland East Cushitic languages also show a Jespersen-type development, though of a rather different type to any we have seen thus far. Starting with Oromo, the situation is as follows (cf. Owens 1985, Bader 2006). Affirmative verbs are fully inflected for person and number in both past and non-past tenses (178)a,b, and the paradigms for both tenses are very similar (though there appears to be more variation in the realization of the vowels of the inflections in the non-past than in the past). Negative non-past verbs are inflected similarly to their affirmative counterparts (albeit apparently without the variability noted for the affirmative non-past) and are marked negative by prefixing *hin* plus high tone on the root syllable of the verb. Negative past verbs, by contrast, take the same prefix *hin* as well as high tone on the root, but do not inflect for person and number, adding instead an invariable *-ne* suffix to the root.

(178) (a) dēnte ~ dēnti ~ dēntu

go.NONPST.3FSG

‘She is going.’

(b) *dēnte*

go.PST.3FSG

‘She went.’

(c) *hin-démtu*

NEG-go.NONPST.3FSG

‘She isn’t going.’

(d) *hin-démne*

NEG-go.PST

‘I/you/he/she/we etc. didn’t go.’ (Oromo; Owens 1985: 66)

The situation in Somali is similar. Affirmative past and non-past verbs are fully (and similarly) inflected, as is the negative non-past verb together with the preverbal negator *má*, while the negative past is formed with the same negator and by the addition to the bare infinitive form (*súgi* in the case of the verb in (179)) of an invariable suffix *-n* (where the infinitive ends in /n/ the past negative form of the verb is identical to the infinitive form).

(179) (a) *sugtaan*

wait.for.PRES.2PL

‘You wait for (it).’

(b) *má-sugtàn*

NEG-wait.for.PRES.2PL

‘You don’t wait for (it).’

(c) sugteen

wait.for.PST.2PL

‘You waited for (it).’

(d) má-sugín

NEG-wait.for.PST

‘I/you/he/she/we etc. didn’t wait for (it).’ (Somali; Saeed 1999: 86-88)

This rather unusual lack of inflection in the negative past of these languages becomes clearer if we consider a third Lowland East Cushitic language, Afar, which appears to have been more conservative in its expression of negation than Somali and Oromo. In Afar, precisely the same situation obtains with respect to the non-past and to the past affirmative. The negative past, however, is clearly composed of the negator *ma*, plus the infinitive form of the verb, plus a clitic auxiliary *-inna* which is fully (but irregularly) inflected for number and person (Bliese 1981: 85). Note that Cushitic languages are largely head-final, with Verb-Aux order as standard. Moreover, the negative copula in Afar is expressed with a form *hinna*, which is inflected identically to the negative past auxiliary. This is shown in Table 6.1 (after Bliese 1981: 85, 111-2).

Table 6.1 Inflection of negative past tense and negative copula in Afar

	<i>m-aggaf-inna</i>	<i>hinna</i>
	'didn't kill'	'is not'
1sg	m-aggaf-inniyo	hinniyo
2sg	m-aggaf-innito	hinnito
3sg	m-aggaf-inna	henna
1pl	m-aggaf-innino	hinnino
2pl	m-aggaf-innitōnu	hinnitōnu
3pl	m-aggaf-innōnu	hinnōnu

If we are right to assume that Afar more closely resembles the situation in Common Lowland East Cushitic (CLEC) than do Oromo and Somali, then the picture is relatively clear. Pre-CLEC negated both past and non-past matrix clause verbs simply by prefixing the negator **ma* (itself derived from an earlier interrogative pronoun; see section 6.1). In CLEC itself the negative copula *hinna* was grammaticalized as a negative past tense auxiliary, following a common grammaticalization path (copula > 'be'-perfect auxiliary), albeit one that is more usually neutral with regard to polarity. As expected, the inflected auxiliary selects the bare infinitive form of the main verb. In Afar, the CLEC situation is preserved, except that the auxiliary becomes a clitic on the verb and the initial /h/ is lost. In Oromo and Somali this clitic presumably first became an inflection and then underwent further phonological erosion, such that it lost its inflection and became a frozen form *-ne* and *-n* respectively. It appears that Pre-Oromo underwent an additional related development such that a bare form of the copula/auxiliary was then

regrammaticalized as a new preverbal negator *hin*, in place of the original CLEC negator *ma*.

To summarize, the quasi-bipartite past tense negative construction of Oromo and Somali appears to be the result of the grammaticalization (and concomitant heavy phonological reduction) of a negative copula, first as a negative past tense auxiliary, and then as an uninflected past tense negative morpheme. A slightly less reduced form of this morpheme then appears to have ousted *ma* as the preverbal negator in all tenses in Oromo. Thus we have here, as in Ethiopian Semitic, another case of what appears to be an unusual set of developments in the syntax of negation, which can however be explained relatively straightforwardly using well-established principles of grammatical change.

6.3 Developments in indefinites

We end this chapter by considering developments in the indefinite pronoun systems of other Afro-Asiatic languages in the light of our discussion of Arabic in chapter 5. Here a lack of data, especially of former states of languages, is a particular problem. Among the languages for which we do have extensive attestation, however, one striking feature is that neither Akkadian nor Egyptian-Coptic ever develop systems of negative concord throughout the entire (very lengthy) period of their attestation. Ancient Egyptian lacks dedicated indefinite pronouns distinct from the words for ‘man’ and ‘thing’ (Haspelmath 1997: 324, citing Gardiner 1957), while Coptic developed a polarity-neutral ‘person’-pronoun *laau* ‘someone, anyone’ but no NPI pronouns or negative quantifiers (Reintges 2004). Akkadian *mimma* ‘something, anything, everything’ and *mamman* ‘someone, anyone’ are attested, with minor

variations, from Old Akkadian (third millennium BCE) to Neo-Babylonian/Neo-Assyrian (first millennium BCE), and at no stage is there any evidence to suggest a development into negative quantifiers, despite frequent co-occurrence with negation (Black, George & Postgate 2000; cf. section 6.1). Thus, while negative concord is undoubtedly common in the world's languages (cf. Haspelmath 1997, Kahrel 1996), the functional pressure on languages which lack it to then develop it is clearly far from irresistible.

By contrast, Hebrew, which will be the main focus of this section, shows significant developments in its indefinite system. Biblical Hebrew did not have negative concord, and appears only to have had one item that was restricted to nonveridical contexts, the 'thing'-pronoun *məʔuma* mentioned in section 6.1. Possible etymologies for this item suggested by Brown, Driver and Briggs (1999: s.v. *məʔuma*) are *ma u-ma* 'what and what' (cf. Latin *quidquid* 'anything'), which is favoured by Faber (1988), and a feminine form of *məʔum* 'blemish, speck, particle' used as a minimizer, which Faber (1988) dismisses. Faber's scepticism is likely to be correct, given that minimizers tend to be restricted just to negative contexts, while *məʔuma* is found both in the context of negation (180)a as well as other nonveridical contexts such as questions (180)b. It is not found in affirmative declarative sentences.

(180) (a) wa-ha-mmeθim ʔen-om yodʕim **məʔuma**
 and-the-dead.PL not-they knowing.PL anything

'But the dead do not know anything.'

(Eccl 9: 5)

have functioned as a minimizer: the name of a given entity viewed as the minimal quantity of that entity (cf. Akkadian *mimma šumšu* ‘anything at all’, lit. ‘anything, it’s name’). Both *klum* and *šum* are found (almost) exclusively in the context of negation (are therefore strong NPIs), as expected in the case of *šum*, given its derivation from a minimizer. There is, however, no evidence to suggest that these items were n-words (according to the definition in (103) in chapter 5) in the pre-revival period.

Modern Israeli Hebrew has a number of n-words, however. These are: the determiners *af*⁹² (183) and *šum* (184), the ‘thing’-pronouns *klum* and *šum davar* (185), the ‘person’-pronoun *af exad* (186), the ‘time’-adverb *af paam* (187), and the ‘place’-adverb *be-šum makom*.

(183) (a) lo raʔiti af xatul

NEG see.PRF.1SG af cat

‘I didn’t see a single cat.’

(b) hayu harbe morim aval af mora

be.PRF.3PL many teacher.MPL but af teacher.FSG

‘There were many male teachers, but no female teacher(s).’

(184) (a) lo raʔiti šum xatul

NEG see.PRF.1SG šum cat

‘I didn’t see any cat(s).’

(b) šamanu neʔumim aval šum xidušim

hear.PRF.1PL speeches but šum new ideas

‘We heard speeches but no new ideas.’

⁹² *Af* is derived from a homophonous Biblical Hebrew focus particle meaning ‘also, even’ (cf. Moroccan Arabic *ħatta* ‘even, a single, no’).

(185) (a) al tikne klum / šum davar

NEG buy.FUT.2MSG n.thing

‘Don’t buy anything.’

(b) A: ma kara

what happen.PRF.3MSG

‘What happened?’

B: klum / šum davar

‘Nothing.’

(186) (a) lo raʔiti af exad

NEG see.PRF.1SG n.one

‘I didn’t see anyone.’

(b) A: mi ba

who come.PRF.3MSG

‘Who came?’

B: af exad

‘No one.’

(187) (a) ani af paʔam lo hayiti šam

I (n)ever NEG be.PRF.1SG there

‘I’ve never been there.’

(b) A: hayita šam

be.PRF.2MSG there

‘Have you been there?’

B: af paʔam

‘Never.’

(188) (a) lo hayiti **be-šum** makom

NEG be.PRF.1SG in-šum place

'I haven't been anywhere.'

(b) A: eyfo hayita

where be.PRF.2MSG

'Where have you been?'

B: **be-šum** makom

in-šum place

'Nowhere.'

(Modern Hebrew; Glinert 1982: 434, 450, 454)

Unlike some of the dialectal Arabic n-words we looked at in chapter 5, all of these n-words are restricted to negative contexts. As such, it is generally assumed that they are negative quantifiers and Modern Hebrew is a negative concord language (e.g. Tonciulescu 2007).

A dissenting voice is Glinert (1982), who analyses these n-words as strong NPIs. His arguments are rather weak, however, amounting to a) that if the n-words were in fact negative quantifiers one would expect a double negation reading to be available in sentences such as (185)a whereas it is not, and b) that it ought to be possible to omit the sentential negator in sentences with a negative quantifier whereas, again, it is not. It is not clear that either of these are valid assumptions. On the other hand, evidence in favour of a negative quantifier analysis of Modern Hebrew n-words comes from a similar range of facts as observed for Maltese in section 5.2: Hebrew

n-words can appear in affirmative sentences in positions where they are not arguments of the verbs, where they nevertheless retain their negative meaning:

(189) ve-hi axšav boxa al klum
and-she now crying.FSG about n.thing

‘And now she’s crying about nothing.’

(<http://www.modiinnews.co.il/blogs/author/ravit/>)

(190) ze adif me-šum davar
this better than-n.thing

‘This is better than nothing.’

(191) adam exad adif me-af exad
person one better than-n.one

‘One person is better than no one.’

(Modern Hebrew)

It appears, then, that Modern Hebrew n-words really are negative quantifiers and that Modern Hebrew is a negative concord language. Tracking this development with respect to the non-negative-concord system of earlier Hebrew is in fact made rather easy by the fact that Hebrew did not exist as a spoken language from the early part of the first millennium until its revival in the late nineteenth century. Those who learnt to speak it in the latter period were necessarily therefore native speakers of other languages, predominantly Yiddish (Zuckermann 2009: 43). On this basis, given that a) negative concord appears not to have been a feature of pre-revival literary Hebrew, b) it does appear, by contrast, to have been a feature of Modern Hebrew

since the earliest days of its revival as a native language (Chanoch 1930: 71-2), and c) Yiddish has long been a negative concord language (see for (192) some examples), this looks like a clear case of syntactic imposition.

(192) (a) **keyner** hot **nit** gezen ire trenn
 nobody AUX.PST.3SG NEG see.PTCP her tears

‘No one saw her tears.’

(b) du vest dayn tsil **keyn** mol **nit** dergreykhn
 you AUX.FUT.2SG your goal no time NEG reach.INF

‘You will never reach your goal.’

(c) zi hot geboyren a kind on **keynems** hilf
 she AUX.PST.3SG bear.PTCP a child with nobody.GEN help

‘She bore a child without anyone’s help.’

(Lockwood 1995: 130-1)

Native Yiddish speakers must have found evidence in the strict co-occurrence with negation of Hebrew indefinites such as *klum* and *šum* for the negative concord system familiar to them from their native language, and interpreted these indefinites as negative quantifiers rather than strong NPIs.

The Modern Hebrew indefinite system is summarized in Table 6.2 (see tables in section 5.2 for legend).

Table 6.2 Indefinites in Modern Hebrew

	negative	n-word	strong	weak
	quantifier		NPI	NPI
Determiner	šum	← šum	(kol) ^a	kol
	af	← af		
Thing	klum	← klum	—	davar
	šum davar	← šum davar		
Person	af exad	← af exad	—	iš
Extent adverb	—	—	—	klal
Time	af paʔam	← af paʔam	leʔolam	—
still/yet	—	(adayin?) ^b	—	—
Place	be-šum makom	← be-šum makom	—	—

^a Glinert (1982) argues for two homophonous determiners *kol*, one a strong NPI, the other restricted to the remaining NPI contexts, on the basis of their different selectional properties.

^b Rosén (1977: 228) claims that *adayin* 'still, yet' is an n-word. Glinert (1982) denies this.

We finish this section with a brief word on indefinites in Aramaic and then Ethiopian Semitic.

Old and Middle Aramaic, like pre-revival Hebrew, seem to have lacked n-words. In Northeastern Neo-Aramaic varieties, however, intense and prolonged contact with Kurdish has resulted in the transfer of the determiner *čī* and the 'thing'-pronoun *hič* from the latter, both of which are n-words in both Kurdish and Neo-Aramaic (Khan 1999: 237, Soane 1913: 30). The question of whether these items are best analysed as strong NPIs or negative quantifiers, however, must for the moment remain a task for future research.

Turning to Ethiopian Semitic, there is no language from this branch for which we can track significant developments in its indefinite system in written texts. However, in Ge'ez we have an early snapshot of the grammar of one Ethiopian Semitic language before it died out as a spoken variety some time before the end of the first millennium (Gragg 1997: 243). Here we have a straightforward negative concord system, whereby clearly negative indefinites *?i-mənt(-hi/-ni)* 'nothing' and *?i-männu(-hi/-ni)* 'no one' (morphologically negated with the general-purpose negator *?i*) co-occur with sentential negation:

(193) wä-?i-tä?ämməxu wä-?i-männa-hi ba-fənot

and-NEG-greet.IMP.2MPL and-NEG-what.ACC-PRT in-way

'and don't greet anyone on the way.'

(Ge'ez, Luke 10: 4; Tropper 2002: 148)

Tigre and Tigrinya, the modern Ethiopian Semitic languages most closely related to Ge'ez, do not appear to be able to form morphologically negative indefinites in this way. Leslau (1945) provides some data on the indefinite pronoun system of Tigre. It does not appear to have any negative quantifiers. Leslau does describe an indefinite 'thing'-pronoun *sema*, which may or may not be an n-word. He glosses it as "(with negation) nothing" and gives the following examples:

(194) (a) **sema**-ma ʔi-räkäbko
sema-EMPH NEG-find.PRF.1SG

‘I haven’t found anything.’

(b) **sema** əmbäl dəhan
sema except good

‘nothing but good.’

(Tigre; Leslau 1945: 192)

Tigrinya does not in general appear to have n-word indefinites (cf. example (173)), with the exception of the negative determiner *wala*, borrowed from Arabic (see section 5.2):

(195) (a) A: ḥadä säb rəʔaka-do
 one man see.PRF.2MSG-Q

‘Did you see anyone?’

B: **wala** ḥadä
 not.even one

‘No one.’

(b) **wala** ḥadä melsi zə-fälliṭ **yällo-n**
 not.even one answer REL-know.IMPF.3MSG NEG.be.3MSG-NEG

‘There’s nobody who knows the answer.’

(Tigrinya; Bernini 2003: 94-5)

Amharic, whose indefinite system is better described than that of any other Ethiopian Semitic language thanks to Leslau's (1995) comprehensive grammar, also lacks morphologically negative indefinites of the type found in Ge'ez. It does, however, have at least one n-word, the 'person'-pronoun *mannəmm*:

(196) (a) balläfäw sammənt bet-ä krəstiyan **mannəmm**
 in.last week house-CONSTR Christian n.one
 al-hedä-mm
 NEG-go.PRF.3MSG-NEG
 'No one went to church last week.'

(b) A: yəhən man näggärä-h
 this who say.PRF.3MSG-you
 'Who told you this?'

B: **mannəmm**

'No one.'

(Amharic; Leslau 1995: 122-3)

This item *mannəmm* is used non-negatively in questions and conditionals and as a free-choice item, as well as in the standard of comparison (197), suggesting that it is not to be analysed as a negative quantifier and that Amharic therefore lacks negative concord.

(197) amarəñña kä-mannəmm yəbäliṭ ənnaggärallä^wh

Amharic than-n.one more speak.CMPR.IMPF.1SG

'I speak Amharic better than anyone.'

(Amharic; Leslau 1995: 121)

Thus we have an interesting situation among the Ethiopian Semitic languages that is the inverse of what has been the case in Europe. Like Latin, Ge'ez is a high language, is no one's native language and was for a long time the only or principal language of literature and liturgy. Unlike Latin, however, it has a clear negative concord system. The fact that a number of languages in Europe lack negative concord, particularly in the case of standard languages, is thought to be typologically somewhat unusual and is therefore often attributed to the influence of Latin (e.g. Haspelmath 1997: 205, 220). But the vernacular Ethiopian Semitic languages we have looked at here appear not to have developed negative concord despite both the hypothetical functional pressure to do so and potential influence from the high variety Ge'ez. Together with the failure of both Akkadian and Egyptian-Coptic to develop negative concord over the course of several millennia, the evidence presented from Afro-Asiatic must therefore cast some doubt on the strength of the functional pressures that have been proposed to account for the prevalence of negative concord crosslinguistically.

For example, Haspelmath (1997: 193-234) invokes two relevant universal preferences: i) for negation to be marked on the verb in ordinary sentence negation (assuming Jespersen's view of negation as predicate denial and Haiman's 1980 principle of form-meaning isomorphism), and ii) Jespersen's (1917) and Horn's

(1989) Negative First Principle, which states that the semantic importance of negation is sufficiently great that it needs to be expressed as early as possible in a sentence. The first of these preferences is upheld in all the languages investigated in this section. The same cannot be said for the second. In fact, we have four languages with SOV order – Akkadian, Tigre, Tigrinya and Amharic – all of which appear to lack negative indefinites (with the exception of Tigrinya's borrowed negative determiner *wala*) despite the fact that sentential negation in these languages routinely follows subjects and objects. Ge'ez, on the other hand, has developed negative indefinites despite their not being necessary to satisfy the Negative First Principle, which is automatically satisfied by its VSO basic word order and negation being a proclitic on the verb.

Of course, proposed functional-typological universals are not invalidated by individual exceptions, but the evidence presented here should give proponents of the Negative First Principle pause for thought. First, verb-final Akkadian did not develop a means of adhering to the Negative First Principle at any stage during the more than two millennia of its recorded history. Second, other than Ge'ez, for which the Negative First Principle appears to be irrelevant in its development of negative concord, all the clear cases of Semitic languages that have developed negative concord (Maltese, Hebrew, Neo-Aramaic) have only done so after intensive contact with other (Indo-European) negative concord languages. Finally, the modern Ethiopian Semitic languages reviewed here, like Akkadian, all violate the Negative First Principle by having no negative indefinites and negation as a clitic on the verb in final position, and this despite the influence one might expect from the high variety, Ge'ez.

Thus we see that consideration of the diachrony of non-European languages with lengthy historical attestation can lead to conclusions which may be surprising from a narrow European perspective. The aim of this chapter has been to supplement the detailed existing work on negation and negative concord in European languages with analyses of similar data in lesser-studied languages, thereby shedding further light on the pathways of change in this domain, and deepening our understanding of the processes and pressures underlying these changes.

7 Conclusion

This thesis has given an account of the historical processes underlying the occurrence of Jespersen's Cycle and other developments in the expression of sentential negation in a number of Afro-Asiatic languages. We have also addressed the issue of negative concord and the question of whether any languages from this family can be said to have developed this property during their recorded history.

We have seen that Jespersen's Cycle is a feature of the histories of Coptic, Modern South Arabian, certain Arabic varieties and certain Berber varieties. I argued in chapter 2 that these developments in all of these languages should be seen as linked by language contact, and that the widespread perception that Jespersen's Cycle is in general too common crosslinguistically to merit an account that postulates contact is likely to be mistaken. Specifically, I argued that the distribution of a bipartite negative construction among the Arabic dialects (across north Africa and into the southwestern Levant and in another region in the south of the Arabian Peninsula, but not elsewhere) is best explained as having been triggered once in Egyptian Arabic following contact with Coptic, and a second time in the south of the Arabian Peninsula following contact with Modern South Arabian, both of which we saw also had a bipartite construction. I also argued that the presence in a number of Berber varieties of a bipartite construction structurally very similar to that of Arabic is best explained as having been triggered by contact with the latter. I supported these arguments with linguistic, historical and dialect-geographical evidence, as well as with the small amount of textual evidence that is available for earlier varieties of vernacular Arabic and Berber.

In chapters 3 and 4 I then developed a model of contact-induced grammatical change that allowed me to give a more explicit account than would otherwise have been possible of how it was that language contact contributed to the development of a bipartite negative construction in these languages. In so doing a wider aim was to contribute to an ongoing change of direction in historical syntactic research whereby the explanation for a given change is assessed on its own merits and not assumed wherever possible to have been the result of purely internal processes. I tried to facilitate this process both by providing a framework for accounting for contact-induced change in terms that approach the explicitness of the tools available to us for accounting for purely internal change, as well as by showing that language contact can simply be thought of as one more source of changes to the primary linguistic data on the basis of which children abduce the grammars of their native languages.

In chapter 5 I focussed more closely on developments in the expression of negation in vernacular Arabic. I gave an account of the restrictions on the stage II construction that we observe today in terms of the emphatic value that this construction would have had when it still competed with the stage I construction in unmarked contexts; and I argued that logical negation should be seen as a property of whole propositions rather than individual lexical items, thus explaining why it is that the two negative expressions *mā* and *-š* can each individually identify a sentence as negative, but when they co-occur in a sentence a double negative (affirmative) interpretation does not result. I suggested that negative concord, where the marker of sentential negation co-occurs with a negative quantifier and the interpretation again contains a single logical negation, should be understood in the same way, though I argued that the only Arabic variety that is straightforwardly describable as a negative

concord language is Maltese, despite n-words being common in other varieties. I also gave an account of the limited distribution of stage III negation in Palestinian, in which I argued that the data cannot be explained by means of the synchronic phonological deletion of preverbal *mā*, but that its earlier deletion in a highly restricted context led to a syntactic reanalysis and the subsequent spread of stage III negation to further syntactically (not phonologically) defined contexts.

Finally, in chapter 6 I used the principles established in previous chapters to look at several similar developments in other Afro-Asiatic languages. We saw that several languages from a number of different branches of Ethiopian Semitic have developed a bipartite negative construction in which the second element has its rather unusual origin in an enclitic conjunction. I argued that this was best explained (following Hetzron 1972) as the result of contact with the Cushitic language Awngi in which there is a bipartite negative construction whose second element is homophonous with an enclitic conjunction. We also looked briefly at negation in a number of Lowland East Cushitic languages, for which I argued on comparative grounds that the uninflected suffix that is characteristic of the negative past tense in several of these languages is the end point of the grammaticalization of a negative copula as a negative past tense ('be'-perfect) auxiliary. Finally, we looked at the development of indefinites in the scope of negation in Afro-Asiatic languages other than Arabic, noting primarily the conspicuous absence of negative concord in many of these languages, even throughout the several millennia of the recorded history of Akkadian and Egyptian-Coptic. We also saw that the languages which clearly have developed negative concord – Hebrew and Aramaic, as well as Maltese as discussed in chapter

5 – seem only to have done so as a result of contact with other languages that had this property.

The historical development of negation has attracted a good deal of interest in previous literature, but this has focussed predominantly on individual European languages in isolation. It is hoped that the focus in this thesis on a range of Afro-Asiatic languages has contributed not only to our understanding of the history of these languages, but also to a more balanced picture of the ways in which the syntactic expression of negation can develop in natural language generally. Of course, in common with all historical studies that focus on developments in a particular syntactic domain, it is also hoped that the discussion in this thesis has contributed to our understanding of the mechanisms that underlie change in grammars more generally.

Works cited

I John = The First Epistle of John. In G. Horner (ed). 1911. *The Coptic version of the New Testament in the Southern dialect: Otherwise called Sahidic and Thebaic*. Oxford: The Clarendon Press.

2MAB = Basil of Caesarea's Second Homily on St Michael Archangel. In L. Depuydt (ed). 1991. *Homiletica from the Pierpont Morgan Library: Seven Coptic Homilies Attributed to Basil the Great, John Chrysostom, and Euodius of Rome*. Leuven: Peeters.

BCJ = John Chrysostom's Encomium on the Bodiless Creatures. In Depuydt (1991).

C = The Martyrdom of St Coluthus. In E. Reymond & J. Barns (eds). 1973. *Four Martyrdoms from the Pierpont Morgan Coptic Codices*. Oxford: Clarendon Press.

Eccl = The Book of Ecclesiastes. In K. Elliger, R. Kittel & W. Rudolph (eds.). 1997. *Biblia Hebraica Stuttgartensia*. Stuttgart: Deutsche Bibelgesellschaft.

Gal = The Epistle of Paul to the Galatians. In Horner (1911).

M.B = Basil of Caesarea's Encomium on St Mercurius. In Depuydt (1991).

MA.B = Basil of Caesarea's First Homily on St Michael Archangel. In Depuydt (1991).

MA.J = John Chrysostom's Homily on St Michael Archangel. In Depuydt (1991).

Num = The Book of Numbers. In Elliger et al. (1997).

Philippians = Paul's Epistle to the Philippians. In Carmel Sant (ed.). 1984. *Il-Bibbja*. Floriana: The Malta Bible Society.

PT = Paese and Thecla. In Reymond & Barnes (1973).

Qur'an = Ali, Ahmed (ed). 1988. *Al-Qur'ān*. Princeton, NJ: Princeton University Press.

Stauffer 1975 = Stauffer, R. (ed.). 1975. *Charles Darwin's Natural Selection; being the second part of his big species book written from 1856 to 1858*. Cambridge: Cambridge University Press.

RA.J = John Chrysostom's Homily on the Resurrection and the Apostles. In Depuydt (1991).

SB = Shenoufe and his Brethren. In Reymond & Barnes (1973).

List of references

- Abraham, Werner. 2003. 'Autonomous and non-autonomous components of 'grammatic(al)ization': economy criteria in the emergence of German negation'. *Sprachtypologie und Universalienforschung* 56: 325-65.
- Abu El-Haija, Lutfi. 1981. 'The acquisition of the negation system in Arabic as spoken in Jordan'. PhD dissertation, the Pennsylvania State University.
- Abu Haidar, Farida. 1979. *A Study of the Spoken Arabic of Baskinta*. Leiden: Brill.
- Abun-Nasr, Jamil. 1987. *A History of the Maghrib*. Cambridge: Cambridge University Press.
- Adila, Aziz. 1996. 'La négation en arabe marocain'. In Chaker & Caubet (1996), 99-116.
- Albright, Adam. 2008. 'Explaining universal tendencies and language particulars in analogical change'. In J. Good (ed.), *Linguistic Universals and Language Change*. Oxford: Oxford University Press, 144-184.
- Altenberg, Evelyn. 1991. 'Assessing first language vulnerability to attrition'. In H. Seliger & R. Vago (eds.), *First Language Attrition*. Cambridge: Cambridge University Press, 189-206.
- & R. Vago. 2004. 'The role of grammaticality judgments in investigating first language attrition'. In M. Schmid, B. Köpke, M. Keijzer & L. Weilemar (eds.), *First Language Attrition: Interdisciplinary Perspectives on Methodological Issues*. Amsterdam: Benjamins, 105-129.
- Andersen, Henning. 1973. 'Abductive and deductive change'. *Language* 49: 765-793.

- As, Arvid. 1962. 'The recovery of a forgotten language through hypnotic age regression: a case report'. *The American Journal of Clinical Hypnosis* 5: 24-29.
- Ashby, William. 1991. 'When does variation indicate linguistic change in progress?'. *Journal of French Language Studies* 1: 1-19.
- Backus, Ad. 2005. 'Codeswitching and language change: one thing leads to another?'. *International Journal of Bilingualism* 9.3-4: 307-340.
- Bader, Christian. 2006. *Parlons oromo*. Paris: L'Harmattan.
- Basset, André. 1952. *La langue berbère*. London: Oxford University Press.
- Bauer, Leonhard. 1913. *Das palästinische Arabisch: die Dialekte des Städters und des Fellachen. Grammatik, Übungen und Chrestomathie*. Leipzig: J.C. Hinrichs.
- Behnstedt, Peter. 1985. *Die nordjemenitischen Dialekte*, vol. 1. Wiesbaden: L. Reichert.
- Benmamoun, Elabbas. 1997. 'Licensing of negative polarity items in Moroccan Arabic'. *Natural Language and Linguistic Theory* 15.2: 263-287.
- 2000. *The Feature Structure of Functional Categories: A Comparative Study of Arabic Dialects*. Oxford: Oxford University Press.
- 2006. 'Licensing configurations: the puzzle of head negative polarity items', *Linguistic Inquiry* 37.1: 141-147.
- Bernini, Giuliano. 2003. 'Learner varieties and language types. The case of indefinite pronouns in non-native Italian'. in A. Giacalone Ramat (ed.), *Typology and Second Language Acquisition*. Berlin: Mouton de Gruyter, 83-124.
- & P. Ramat. 1992. *La frase negativa nelle lingue d'Europa*. Bologna: Il Mulino.

- & P. Ramat. 1996. *Negative Sentences in the Languages of Europe: A Typological Approach*. Berlin: Mouton de Gruyter.
- Berwick, Robert. 1985. *The Acquisition of Syntactic Knowledge*. Cambridge, Mass.: MIT Press.
- Besten, Hans den. 1986. 'Double negation and the genesis of Afrikaans'. In P. Muysken & N. Smith (eds.), *Substrata versus Universals in Creole Languages. Papers from the Amsterdam Creole Workshop, April 1985*. Amsterdam: Benjamins, 185-230.
- Beyer, Klaus. 2009. 'Double negation marking: a case of contact-induced grammaticalization in West Africa?'. In N. Cyffer, E. Ebermann & G. Ziegelmeier (eds.), *Negation Patterns in West African Languages and Beyond*. Amsterdam: Benjamins, 205-221.
- Biberauer, Theresa. 2008. 'Doubling and omission: insights from Afrikaans negation'. In S. Barbiers, M. van der Ham, O. Koenen & M. Lekakou (eds.), *Microvariations in Syntactic Doubling*. Bingley: Emerald, 103-140.
- Bishai, Wilson. 1962. 'Coptic grammatical influence on Egyptian Arabic'. *Journal of the American Oriental Society* 82.3: 285-289.
- Black, Jeremy, A. George, & N. Postgate. 2000. *A Concise Dictionary of Akkadian*. Wiesbaden: Harrassowitz.
- Blau, Joshua. 1960. *Syntax des palästinensichen Bauerndialektes von Bīr-Zēt, auf Grund der "Volkserzählungen aus Palästina" von Hans Schmidt und Paul Kahle*. Walldorf, Hessen: Verlag für Orientkunde Dr. H. Vorndran.
- 1965. *The Emergence and Linguistic Background of Judaeo-Arabic*. Oxford: Oxford University Press.

- Bley-Vroman, Robert. 1989. 'What is the logical problem of foreign language learning?'. In S. Gass, & J. Schachter (eds.), *Linguistic Perspectives on Second Language Acquisition*. Cambridge: Cambridge University Press, 41-68.
- 2009. 'The evolving context of the Fundamental Difference Hypothesis'. *Studies in Second Language Acquisition* 31: 175-198.
- Bliese, Loren. 1981. *A Generative Grammar of Afar*. Dallas, Texas: Summer Institute of Linguistics.
- Blom, Elma, D. Polišenská & F. Weerman. 2006. 'Effects of age on the acquisition of agreement inflection'. *Morphology* 16.2: 313-336.
- Borg, Albert & M. Azzopardi-Alexander. 1997. *Maltese*. London: Routledge.
- Borg, Alexander. 2004. *A Glossary of Cypriot Maronite Arabic (Arabic-English)*. Leiden: Brill.
- Borsley, Robert, M. Tallerman & D. Willis. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Bosch, Elena, F. Calafell, A. Pérez-Lezaun, J. Clarimón, D. Comas, E. Mateu, R. Martínez-Arias, B. Morera, Z. Brakez, O. Akhayat, A. Sefiani, G. Hariti, A. Cambon-Thomsen & J. Bertranpetit. 2000. 'Genetic structure of north-west Africa revealed by STR analysis'. *European Journal of Human Genetics* 8.360-366.
- Bosch, Laura & N. Sebastián-Gallés. 2001. 'Early language differentiation in bilingual infants'. In J. Cenoz & F. Genesee (eds.), *Trends in Bilingual Acquisition*. Amsterdam: Benjamins, 71-94.
- Boumalk, Abdallah. 1996. 'La négation en berbère marocain'. In Chaker & Caubet (1996), 35-48.

- Bowerman, Melissa. 1988. 'The 'no negative evidence' problem: How do children avoid constructing an overly general grammar?'. In J. Hawkins (ed.), *Explaining Language Universals*. Oxford: Blackwell, 73-101.
- Breitbarth, Anne. 2009. 'The development of negation in Dutch and Low German'. In Breitbarth et al. (in preparation).
- & L. Haegeman. 2008. 'Not continuity, but change: stable stage II in Jespersen's cycle'. Paper presented at the Annual Meeting of the LAGB, University of Essex, Colchester.
- , C. Lucas & D. Willis. Forthcoming. 'Incipient Jespersen's Cycle: the (non-)grammaticalization of new negative markers'. In J. Fleischer & H. Simon (eds.), *Comparing Diachronies*. Tübingen: Max Niemeyer.
- , C. Lucas & D. Willis (eds.) In preparation. *The History of Negation in the Languages of Europe and the Mediterranean, vol. 2: Case Studies*. Oxford: Oxford University Press.
- Brincat, Joseph. 1995. *Malta 870 - 1054: Al-Ḥimyarī's Account and its Linguistic Implications*. Valetta: Said International.
- Brown, Francis, S. Driver & C. Briggs. 1999. *The Brown-Driver-Briggs Hebrew and English Lexicon: With an Appendix Containing the Biblical Aramaic*. Peabody, Mass.: Hendrickson Publishers.
- Brugnatelli, Vermondo. 1986. 'La negazione discontinua in berbero e in arabomagrebino', in G. Bernini & V. Brugnatelli (eds.), *Atti della 4a Giornata di studi camito-semitici ed indoeuropei*, Milan: Unicopli, 53-62.

- Brustad, Kristen. 2000. *The Syntax of Spoken Arabic: A Comparative Study of Moroccan, Egyptian, Syrian and Kuwaiti Dialects*. Washington, DC: Georgetown University Press.
- Burrdige, Kate. 1993. *Syntactic Change in Germanic: Aspects of Language Change in Germanic, with Particular Reference to Middle Dutch*. Amsterdam: Benjamins.
- Bybee, Joan. 2003. 'Mechanisms of change in grammaticization: the role of frequency', in R. Janda and B. Joseph (eds.), *Handbook of Historical Linguistics*. Oxford: Blackwell, 602-623.
- & P. Hopper (eds.) 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins.
- Carruthers, Peter. 2005. 'The case for massively modular models of mind'. In R. Stainton (ed.), *Contemporary Debates in Cognitive Science*. Oxford: Blackwell, 205-225.
- Cassar, Carmel. 2000. *A Concise History of Malta*. Msida: Mireva.
- Catalani, Luigi. 2001. *Die Negation im Mittelfranzösischen*. Frankfurt am Main: Peter Lang.
- Caubet, Dominique. 1993. *L'arabe marocain*. Paris: Peeters.
- 1996. 'La négation en arabe maghrébin'. In Chaker & Caubet (1996), 79-97.
- 2000-1. 'Questionnaire de dialectologie du Maghreb (d'après les travaux de W. Marçais, M. Cohen, G.S. Colin, J. Cantineau, D. Cohen, Ph. Marçais, S. Levy, etc.)'. *Estudios de dialectologica norteafricana y andalusí* 5: 73-92.
- Chaâbane, Nadia, 1996. 'La négation en arabe tunisien'. in Chaker & Caubet (1996), 117-134.

- Chaker, Salem. 1996. 'Remarques préliminaires sur la négation en berbère'. In Chaker & Caubet (eds.), 117-134.
- & D. Caubet (eds.). 1996. *La négation en berbère et en arabe maghrébin*. Paris: L'Harmattan.
- Chanoch, Irene. 1930. *Fremdsprachliche Einflüsse im modernen Hebraisch*. Heidelberg: Winter.
- Chomsky, Noam. 1980. *Rules and Representations*. New York: Columbia University Press.
- 1985. Transcript of an informal interview with Noam Chomsky on the topic of bilingualism, conducted by F. Grosjean, May 29th, 1985.
- 'Minimalist inquiries: the framework'. In R. Martin, D. Michaels & J. Uriagereka (eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, Mass.: MIT Press, 89-156.
- Clahsen, Harald & P. Muysken. 1986. 'The availability of Universal Grammar to adult and child learners: a study of the acquisition of German word order'. *Second Language Research* 2: 93-119.
- 1989. 'The UG paradox in L2 acquisition'. *Second Language Research* 5.1: 1-29.
- Clark, Robin & I. Roberts. 1993. 'A computational approach to language learnability and language change.' *Linguistic Inquiry* 24: 299-345.
- Coetsem, Frans van. 1988. *Loan Phonology and the Two Transfer Types in Language Contact*. Dordrecht: Foris.
- 2000. *A General and Unified Theory of the Transmission Process in Language Contact*. Heidelberg: Winter.

- Cook, Vivian. 2002. 'Background to the L2 user'. In V. Cook (ed.), *Portraits of the L2 User*. Clevedon: Multilingual Matters, 1-28.
- Corriente, Federico. 1977. *A Grammatical Sketch of the Spanish Arabic Dialect Bundle*. Madrid: Instituto Hispano-Árabe de Cultura.
- Cowart, Wayne. 1997. *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. London: SAGE.
- Cowell, Mark. 1964. *A Reference Grammar of Syrian Arabic*. Georgetown: Georgetown University Press.
- Crawford, David. 2002. 'Morocco's invisible Imazighen'. *Journal of North African Studies* 7.1: 53-70.
- Dahl, Östen. 1979. 'Typology of sentence negation'. *Linguistics* 17, 79-106.
- Davies, Humphrey. 1981. 'Seventeenth-century Egyptian Arabic: a profile of the colloquial material in Yūsuf al-Širbīnī's "Hazz al-Quḥūf fī Šarḥ Qaṣīd Abī Šādūf"'. Ph.D. dissertation, University of California, Berkley.
- Davies, Virginia. 1973. *The Syntax of the Negative Particles bw and bn in Late Egyptian*. Münchner Ägyptologische Studien 29. Munich: Deutscher Kunstverlag.
- DeKeyser, Robert. 2000. 'The robustness of critical period effects in second language acquisition'. *Studies in Second Language Acquisition* 22: 499-533.
- Depuydt, Leo. 1991. 'Introduction'. In L. Depuydt (ed.), *Homiletica from the Pierpont Morgan Library: Seven Coptic Homilies Attributed to Basil the Great, John Chrysostom, and Euodius of Rome*. Leuven: Peeters.
- Detges, Ulrich. 2003. 'La grammaticalisation des constructions de négation dans une perspective onomasiologique, ou: la déconstruction d'une illusion d'optique'. In P.

- Koch & A. Blank (eds.), *Kognitive romanische Onomasiologie und Semasiologie*. Tübingen: Niemeyer, 213–33.
- & R. Waltereit. 2002. 'Grammaticalization vs. reanalysis: a semantic-pragmatic account of functional change in grammar'. *Zeitschrift für Sprachwissenschaft* 21.2: 151-195.
- Devos, Maud, & J. van der Auwera. 2009. 'Jespersen Cycles in Bantu: double and triple negation.' Ms., Royal Museum for Central Africa and University of Antwerp.
- Diakonoff, Igor. 1996. 'Some reflections on the Afrasian linguistic macrofamily'. *Journal of Near Eastern Studies* 55: 293-4.
- Donhauser, Karin. 1996. 'Negationssyntax in der deutschen Sprachgeschichte: Grammatikalisierung oder Degrammatikalisierung?'. In E. Lang & G. Zifonun (eds.), *Deutsch – Typologisch*. Berlin: Walter de Gruyter, 201-17.
- 1998. 'Negationssyntax in Althochdeutschen: ein sprachhistorisches Rätsel und der Weg zu seiner Lösung'. In K. Donhauser & L. Eichinger (eds.), *Deutsche Grammatik: Thema in Variationen*. Heidelberg: Universitätsverlag C. Winter, 283-98.
- Döpke, Susanne. 1998. 'Competing language structures: the acquisition of verb placement by bilingual German-English children'. *Journal of Child Language* 25: 558-584.
- Driver, Godfrey. 1925. *A Grammar of the Colloquial Arabic of Syria and Palestine*. London: Probsthain & Co.
- Dryer, Matthew S. 2008. 'Negative morphemes'. In M. Haspelmath, M. Dryer, D. Gil, & B. Comrie (eds.), *The World Atlas of Language Structures Online*. Munich: Max

- Planck Digital Library, chapter 112. Available online at <http://wals.info/feature/112>. Accessed on 2009-09-28.
- Durand, Olivier. 2004. *L'arabo del Marocco: elementi di dialetto standard e mediano*. Rome: Università degli Studi La Sapienza.
- Eades, Diana. 1993. 'Aboriginal English'. In S. Wurm, P. Mühlhäusler & D. Tyron (eds.), *Atlas of Languages of Intercultural Communication in the Pacific, Asia and the Americas*. Berlin: Mouton de Gruyter, 133-142.
- Eckardt, Regina. 2006. *Meaning Change in Grammaticalization: An Inquiry into Semantic Reanalysis*. Oxford: Oxford University Press.
- Eckmann, Fred. 1996. 'A functional-typological approach to second language acquisition theory'. In W. Ritchie & T. Bhatia (eds.), *Handbook of Second Language Acquisition*. New York: Academic Press, 195-212.
- Ehret, Christopher. 1995. *Reconstructing Proto-Afro-Asiatic: Vowels, Tone, Consonants and Vocabulary*. Berkeley: University of California Press.
- Elhalimi, Brahim. 1996. 'La négation dans le parler arabe de Mazouna (Ouest Algérien)'. In Chaker & Caubet (1996), 135-162.
- Elias, David. 2007. 'A new negative morpheme in Tigre'. Paper presented at *NACAL 35*, San Antonio, Texas.
- Espinal, Maria Teresa. 1992. 'Expletive negation and logical absorption.' *The Linguistic Review* 9.4: 333-358.
- Eythórsson, Thórhallur. 2002. 'Negation in C: the syntax of negated verbs in Old Norse'. *Nordic Journal of Linguistics* 25: 190-224.

- Faber, Alice. 1988. 'Indefinite pronouns in early Semitic', in Y. Arbeitman (ed.), *Fucus: A Semitic/Afrasian Gathering in Remembrance of Albert Ehrman*. Amsterdam: Benjamins.
- 1991. 'The diachronic relationship between negative and interrogative markers in Semitic'. In A. Kaye (ed.), *Semitic studies in honor of Wolf Leslau*, vol. 1. Wiesbaden: Otto Harrassowitz, 411-429.
- 1997. 'Genetic subgrouping of the Semitic languages'. In Hetzron (1997a), 3-15.
- Fadhlaoui-Zid, K., S. Plaza, F. Calafell, M. Ben Amor, D. Comas, A. Bennamar El gaaied. 2004. 'Mitochondrial DNA heterogeneity in Tunisian Berbers'. *Annals of Human Genetics* 68: 222-233.
- Fernández, Miguel. 1999. *An Introductory Grammar of Rabbinic Hebrew*. Translated by John Elwolde. Leiden: Brill.
- Flores, Cristina. 2007. 'Sprachverlust im Kontext deutsch-portugiesischer Remigration'. In O. Grossegeesse & C. Flores (eds.), *Wildern in luso-austro-deutschen Sprach- und Textgefilden. Festschrift für Erwin Koller*. Braga: Cehum, Universidade do Minho.
- Fodor, Janet. 1998. 'Unambiguous triggers'. *Linguistic Inquiry* 29.1: 1-36.
- & W. Sakas. 2005. 'The Subset Principle in syntax: costs of compliance'. *Journal of Linguistics* 41: 513-569.
- Footnick, Rosalie. 2007. 'A hidden language: recovery of a 'lost' language is triggered by hypnosis.' In B. Köpke, M. Schmid, M. Keijzer & S. Dostert (eds.), *Language Attrition: Theoretical perspectives*. Amsterdam: Benjamins, 169-187.

- Fromm, Erika. 1970. 'Age regression with unexpected reappearance of a repressed childhood language'. *International Journal of Clinical and Experimental Hypnosis* 18.2: 79-88.
- Fuß, Eric, I. Roberts & C. Trips. 2009. 'Language change and language acquisition: the actuation problem revisited'. Paper presented at the Annual Meeting of the LAGB, University of Edinburgh.
- Gamal-Eldin, Saad. 1967. *A Syntactic Study of Egyptian Colloquial Arabic*. The Hague: Mouton.
- Gardiner, Alan. 1904. 'The word *iwn3*'. *Zeitschrift für Ägyptische Sprache und Altertumskunde* 41: 130-135.
- 1957. *Egyptian Grammar: Being an Introduction to the Study of Hieroglyphs*. 3rd edition. London: Oxford University Press.
- Gary, Judith. & S. Gamal-Eldin. 1981. *Cairene Egyptian Colloquial Arabic*. London: Croom Helm.
- Gast, Volker & J. van der Auwera. 2009. 'What is contact-induced grammaticalization? Evidence from Mayan and Mixe-Zoquean languages'. Ms., Freie Universität Berlin and Universiteit Antwerpen.
- Gawlitzek-Maiwald, Ira, R. Tracy & A. Fritzenschaft. 1992. 'Language acquisition and competing linguistic representations: the child as arbiter'. In J. Meisel (ed.), *The Acquisition of Verb Placement*. Dordrecht: Kluwer, 139-179.
- Gelderen, Elly van. 2008. 'The negative cycle', *Linguistic Typology* 12.2: 195-243.
- Genesee, Fred. 1989. 'Early bilingual development: one language or two?' *Journal of Child Language* 16: 161-179.

- Giannakidou, Anastasia. 1998. *Polarity Sensitivity as (Non)veridical Dependency*. Amsterdam: Benjamins.
- 2000. 'Negative... concord?'. *Natural Language and Linguistic Theory* 18.3: 457-423.
- 2006. 'N-words and Negative Concord'. In M. Everaert and H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*. Oxford: Blackwell, 327-391.
- 2007. 'The landscape of EVEN'. *Natural Language and Linguistic Theory* 25: 39-81.
- Glinert, Lewis. 1982. 'Negative and non-assertive in contemporary Hebrew'. *Bulletin of the School of Oriental and African Studies* 45: 434-470.
- Gragg, Gene. 1997. 'Ge'ez', in Hetzron (1997a), 242-260.
- Greenberg, Joseph. 1963. *The Languages of Africa*. 2nd edition. Bloomington: Indiana University Press.
- 1987. *Language in the Americas*. Palo Alto: Stanford University Press.
- Grice, H. P. 1978. 'Further notes on logic and conversation'. In P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*. New York: Academic Press, 113-128.
- Groll, Sarah. 1970. *The Negative Verbal System of Late Egyptian*. Oxford: Oxford University Press.
- Guichard, Pierre. 2000. *Al-Andalus: 711–1492*. Paris: Hachette Littératures.
- Guy, Gregory. 1990. 'The sociolinguistic types of language change'. *Diachronica* 7.1: 47-67.
- Haegeman, Lilianne & R. Zanuttini. 1991. 'Negative heads and the Neg Criterion'. *The Linguistic Review* 8: 233-252.

- Haiman, John. 1980. 'The iconicity of grammar: isomorphism and motivation', *Language* 56.3: 565-589.
- Håkansson, Gisela, M. Pienemann, & S. Sayehli. 2002. 'Transfer and typological proximity in the context of second language processing'. *Second Language Research* 18.3: 250-273.
- Hale, Mark. 1998. 'Diachronic syntax'. *Syntax* 1.1: 1-18.
- Hansen, Maj-Britt Mosegaard. 2009. 'The negative cycle in French'. In Breitbarth et al. (in preparation).
- Harrell, Richard (ed.). 1966. *A Dictionary of Moroccan Arabic: Arabic-English*. Washington D.C.: Georgetown University Press.
- Harris, Alice & L. Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- Haspelmath, Martin. 1997. *Indefinite Pronouns*. Oxford: Oxford University Press.
- 1999. 'Why is grammaticalization irreversible?'. *Linguistics* 37: 1043-68.
- & Caruana, J. 1996. 'Indefinite pronouns in Maltese', *Rivista di Linguistica* 8.1: 213-227.
- , M. Dryer, D. Gil & B. Comrie, (eds.). 2008. *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library.
- Haugen, Einar. 1950. 'The analysis of linguistic borrowing'. *Language* 26: 210-231.
- Haznedar, Belma. 1995. 'Acquisition of English by a Turkish child: on the development of VP and negation'. Paper presented at Language Acquisition Research Symposium (LARS), University of Utrecht. (Ms., University of Durham).
- Heath, Jeffrey. 2002. *Jewish and Muslim Dialects of Moroccan Arabic*. London and New York: Routledge Curzon.

- 2005. *A Grammar of Tamashek (Tuareg of Mali)*. Berlin: Mouton de Gruyter.
- Heine, Bernd and T. Kuteva. 2003. 'On contact-induced grammaticalization'. *Studies in Language* 27.3: 529-572.
- 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press.
- 2007. 'Identifying instances of contact induced grammatical replication'. Paper presented at Symposium on Language Contact and the Dynamics of Language: Theory and Implications, Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Herburger, Elena. 2001. 'The negative concord puzzle revisited'. *Natural Language Semantics* 9.3: 289-333.
- Hetzron, Robert. 1972. *Ethiopian Semitic: Studies in Classification*. Journal of Semitic Studies Monograph no. 2. Manchester: Manchester University Press.
- (ed.) 1997a. *The Semitic Languages*. London: Routledge.
- 1997b. 'Outer South Ethiopic', in Hetzron (1997a), 535-549.
- Hinds, Martin & E. Badawi. 1986. *A Dictionary of Egyptian Arabic: Arabic-English*. Beirut: Librairie du Liban.
- Hirschbühler, Paul & M. Labelle. 1994. 'Changes in verb position in French negative infinitival clauses'. *Language Variation and Change* 6: 149-178.
- Hoeksema, Jack. 1994. 'On the grammaticalization of negative polarity items'. In S. Gahl, A. Dolbey and C. Johnson (eds.), *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 273-282.

- 1997. 'Negation and negative concord in Middle Dutch'. In D. Forget, P. Hirschbühler, F. Martineau & M. Rivero (eds.), *Negation and Polarity: Syntax and Semantics*. Amsterdam: Benjamins, 139-156.
- Hoffman, Katherine. 2006. 'Berber language ideologies, maintenance, and contraction: gendered variation in the indigenous margins of Morocco'. *Language & Communication* 26.2: 144-167.
- 2008. *We Share Walls: Language, Land and Gender in Berber Morocco*. Oxford: Blackwell.
- Horn, Laurence. 1989. *A Natural History of Negation*. Chicago: Chicago University Press.
- de Houwer, Annick. 1998. 'Comparing error frequencies in monolingual and bilingual acquisition'. (Peer commentary on Müller 1998). *Bilingualism: Language and Cognition* 1.3: 173-174.
- Hoyt, Frederick. 2005. 'Negative Concord in Two Dialects of Arabic'. Ms., University of Texas at Austin.
- 2006. 'An Arabic Wackernagel clitic? The morphosyntax of negation in Palestinian Arabic'. In M. Mughazy (ed.), *Perspectives on Arabic Linguistics XX*. Amsterdam: Benjamins, 105-134.
- Hudson, Grover. 1997. 'Amharic and Argobba', in Hetzron (1997a), 457-485.
- Iaaich, Jamal. 1996. 'La négation en hassaniyya de Tan-Tan (Maroc)'. In Chaker & Caubet (1996), 163-176.
- Ingham, Richard. Forthcoming. 'Negative co-ordination in the history of English'. In A. Breitbarth, C. Lucas, S. Watts & D. Willis (eds.), *Continuity and Change in Grammar*. Amsterdam: Benjamins.

- Israel, Michael. 1998. 'The rhetoric of grammar: scalar reasoning and polarity sensitivity'. PhD dissertation, University of California, San Diego.
- 2001. 'Minimizers, maximizers and the rhetoric of scalar reasoning'. *Journal of Semantics* 18: 297-331.
- Jäger, Agnes. 2008. *History of German negation* (Linguistics today 188). Amsterdam: Benjamins.
- Janse, Mark. 2008. 'Clitic doubling from Ancient to Asia Minor Greek'. In D. Kallulli & L. Tasmovski (eds.), *Clitic Doubling in the Balkan Languages*. Amsterdam: Benjamins, 165-202.
- Jespersen, Otto. 1917. *Negation in English and Other Languages*. Copenhagen: Høst.
- 1924. *The Philosophy of Grammar*. London: George Allen and Unwin.
- Johanson, Lars. 2002. *Structural Factors in Turkic Language Contact*. London: Curzon.
- de Jong, Rudolf. 2000. *A Grammar of the Bedouin Dialects of the Northern Sinai Littoral*. Leiden: Brill.
- Johnson, Jacqueline, K. Shenkman, E. Newport & D. Medin. 1996. 'Indeterminacy in the grammar of adult language learners'. *Journal of Memory and Language* 35: 335-352.
- Johnstone, Thomas. 1977. *Ḥarsūsi Lexicon and English-Ḥarsūsi Word List*. London: Oxford University Press.
- 1981. *Jibbāli lexicon*. London: Oxford University Press.
- Julien, Charles-André. 1961. *Histoire de l'Afrique du Nord*. Paris: Payot.

- Kahrel, Peter. 1996. 'Aspects of Negation'. Ph.D. dissertation, University of Amsterdam.
- Karatsareas, Petros. 2007. 'Contact-induced morphological change: the 'leap' from fusional to inflectional morphology in Cappadocian'. Mphil thesis, University of Cambridge.
- Keenan, Edward. 2002. 'Explaining the creation of reflexive pronouns in English'. In D. Minkova and R. Stockwell (eds.), *Studies in the History of English: A Millennial Perspective*. Berlin: Mouton de Gruyter, 325-355.
- Kemenade, Ans van. 2000. 'Jespersen's Cycle revisited: formal properties of grammaticalization'. In S. Pintzuk, G. Tsoulas & A. Warner (eds.), *Diachronic Syntax: Models and Mechanisms*. Oxford: Oxford University Press, 51-74.
- Khalafallah, Abdelghany. 1969. *A Descriptive Grammar of Suci:di Colloquial Arabic*. The Hague: Mouton.
- Khan, Geoffrey. 1999. *A Grammar of Neo-Aramaic: The Dialect of the Jews of Arbel*. Leiden: Brill.
- Kickasola, Joseph. 1975. 'Sahidic Coptic (*N*)... *AN* negation patterns: a morpho-syntactic description of sentences and adjuncts'. Ph.D. dissertation, Brandeis University.
- King, Ruth. 2000. *The Lexical Basis of Grammatical Borrowing: A Prince Edward Island Case Study*. Amsterdam: Benjamins.
- Kiparsky, Paul. 1996. 'The shift to head-initial VP in Germanic'. In H. Thráinsson, J. Peter, and S. Epstein (eds.), *Studies in Comparative Germanic Syntax*, vol. 2. Dordrecht: Kluwer, 140-179.

- & Condoravdi, Cleo. 2006. 'Tracking Jespersen's Cycle'. In M. Janse, B. Joseph & A. Ralli (eds.), *Proceedings of the 2nd International Conference of Modern Greek Dialects and Linguistic Theory*. Patras: University of Patras, 172-197.
- Kogan, Leonid. 1997. 'Tigrinya'. In Hetzron (1997a), 424-445.
- Köpke, Barbara. 2004. 'Neurolinguistic aspects of attrition'. *Journal of Neurolinguistics* 17: 3-30.
- 2007. 'Language attrition at the crossroads of brain, mind and society'. In B. Köpke, M. Schmid, M. Keijzer & S. Dostert (eds.), *Language Attrition: Theoretical Perspectives*. Amsterdam: Benjamins, 9-37.
- Kraft, Charles & A. Kirk-Greene. 1973. *Hausa*. London: Hodder and Stoughton.
- Kroch, Anthony & A. Taylor. 1997. 'Verb movement in Old and Middle English: dialect variation and language contact'. In A. van Kemenade & N. Vincent (eds.), *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 297-325.
- Kyrris, Costas. 1985. *History of Cyprus*. Nicosia: Nicocles.
- Lafkioui, Mena. 1996. 'La négation en tarifit'. In Chaker & Caubet (1996), 49-77.
- Laka, Itziar. 1990. 'Negation in syntax: on the nature of functional categories and projections'. PhD dissertation, MIT.
- Lambdin, Thomas. 1982. *Introduction to Sahidic Coptic*. Macon, GA: Mercer University Press.
- Lamberti, Marcello. 1991. 'Cushitic and its classifications'. *Anthropos* 86: 552-561.
- Lane. Edward. 1863. *An Arabic-English Lexicon*. London: Williams and Norgate.

- Lasnik, Howard & N. Sobin. 2000. 'The who/whom puzzle: on the preservation of an archaic linguistic feature'. *Natural Language and Linguistic Theory* 18: 343-371.
- Lass, Roger. 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
- Lefebvre, Claire. 2001. 'Relexification in creole genesis and its effects on the development of the creole'. In N. Smith & T. Venestra (eds.), *Creolization and Contact*. Amsterdam: Benjamins, 9-42.
- Lehman, Christian. 1984. *Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: Gunter Narr Verlag.
- Leslau, Wolf. 1945. 'Grammatical sketches in Tigré (North Ethiopic): dialect of Mensa'. *Journal of the American Oriental Society* 65.3: 164-203.
- 1995. *Reference Grammar of Amharic*. Wiesbaden: Harrassowitz Verlag.
- 1999. *Zway Ethiopic Documents: Grammar and Dictionary*. Wiesbaden: Harrassowitz Verlag.
- Lewicki, Tadeusz. 1934. 'Quelques textes inédits en vieux berbère provenant d'une chronique Ibādite anonyme'. *Revue des Etudes Islamiques* 3: 275-296.
- Lightfoot, David. 1989. 'The child's trigger experience: Degree-0 learnability'. *Behavioral and Brain Sciences* 12: 321-75.
- 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge, Mass.: MIT Press.
- 1997. 'Shifting triggers and diachronic reanalyses'. In A. van Kemenade & N. Vincent (eds.), *Parameters of Morphosyntactic Change*. Cambridge: Cambridge University Press, 253-272.

- Lipinski, Edward. 1997. *Semitic Languages: Outline of a Comparative Grammar*.
Leuven: Peeters.
- Lockwood, William. 1995. *Lehrbuch der modernen jiddischen Sprache*. Hamburg:
Buske.
- Longobardi, Giuseppe. 2001. 'Formal syntax, diachronic minimalism, and etymology:
the history of French *chez*'. *Linguistic Inquiry* 32: 275-302.
- Loprieno, Antonio. 1995. *Ancient Egyptian: A Linguistic Introduction*. Cambridge:
Cambridge University Press.
- Lucas, Christopher. 2007. 'Jespersen's Cycle in Arabic and Berber'. *Transactions of
the Philological Society* 105.3: 398-431.
- & E. Lash, Elliott. Forthcoming. 'Contact as catalyst: the case for Coptic
influence in the development of Arabic negation'. *Journal of Linguistics*.
- Ludolf, Hiob. 1698. *Grammatica Linguae Amharicae*. Frankfurt.
- Luttrell, Anthony. 1975. 'Approaches to medieval Malta'. In A. Luttrell (ed.), *Medieval
Malta: Studies on Malta before the Knights*. London: The British School at Rome,
1-70.
- Mad'aj, 'Abd al-Muhsin. 1988. *The Yemen in Early Islam 9-233/630-847: A Political
History*. London: Ithaca Press.
- Maddieson, Ian. 2008. 'Glottalized consonants'. In M. Haspelmath, M. Dryer, D. Gil,
& B. Comrie (eds.), *The World Atlas of Language Structures Online*. Munich: Max
Planck Digital Library, chapter 7. Available online at <http://wals.info/feature/7>.
Accessed on 2009-09-28.
- Mailhammer, Robert. 2007. *The Germanic Strong Verbs: Foundations and
Development of a New System*. Berlin: Mouton de Gruyter.

- Malkiel, Yakov. 1981. 'Drift, slope, and slant: background of, and variations upon, a Sapirian theme'. *Language* 57.3: 535-570.
- Manni, Franz, P. Leonardi, A. Barakat, H. Rouba, E. Heyer, M. Klintschar, K. McElreavey, & L. Quintana-Murci. 2002. 'Y-chromosome analysis in Egypt suggests a genetic regional continuity in north-eastern Africa'. *Human Biology* 74.5: 645-658.
- Martineau, France & R. Mougeon. 2003. 'A sociolinguistic study of the origins of *ne* deletion in European and Quebec French'. *Language* 79: 118-52.
- Matras, Yaron. 1999. 'The state of present-day Domari in Jerusalem'. *Mediterranean Language Review* 11: 1-58.
- 2007. 'Grammatical borrowing in Domari'. In Y. Matras and J. Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter, 151-164.
- & J. Sakel. 2007. 'Investigating the mechanisms of pattern replication in language convergence'. *Studies in Language* 31.4: 829-865.
- Mazzon, Gabriella. 2004. *A History of English Negation*. London: Pearson Longman.
- McCawley, James. 1988. *The Syntactic Phenomena of English*. Chicago: University of Chicago Press.
- Meillet, Antoine. 1912. 'L'évolution des formes grammaticales'. *Scientia* 12: 384-400.
- Meisel, Jürgen. 1989. 'Early differentiation of languages in bilingual children'. In K. Hyltenstam & L. Obler (eds.), *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss*. Cambridge: Cambridge University Press, 13-40.
- 1995. 'Parameters in acquisition'. In P. Fletcher & B. MacWhinney (eds.), *A Handbook of Child Language*. Oxford: Blackwell, 10-35.

- M. 1997. 'The acquisition of the syntax of negation in French and German: contrasting first and second language development'. *Second Language Research* 13.3: 227-263.
- 2001. 'The simultaneous acquisition of two first languages. Early differentiation and subsequent development of grammars'. In J. Cenoz & F. Genesee (eds.), *Trends in Bilingual Acquisition*. Amsterdam: Benjamins, 11-41.
- 2004. 'The bilingual child' in T.K. Bhatia & W.C. Ritchie (eds.), *The Handbook of Bilingualism* (Blackwell Handbooks in Linguistics), Oxford: Blackwell Publishers, 91-113.
- 2008. 'A quest for (transmission) failure: on child bilingualism as the locus of diachronic syntactic change'. Paper presented at *Transmission and Diffusion*, Radboud Universiteit, Nijmegen.
- 2009. 'Second language acquisition in early childhood'. *Zeitschrift für Sprachwissenschaft* 28: 5-34.
- Meltzer, Edmund. 1990. 'The "prehistory" of Late Egyptian *i(w)n(3)*: a hypothesis'. *Göttinger Miszellen* 114: 71-79.
- Mettouchi, Amina. 1996. 'La négation dans les langues du Maghreb: Synthèse'. In Chaker & Caubet (1996), 177-195.
- Mitchell, T. F. 1956. *An Introduction to Egyptian Colloquial Arabic*. Oxford: Oxford University Press.
- Möhring, Anja & J. Meisel. 2003. 'The Verb-Object parameter in simultaneous and successive acquisition of bilingualism'. In N. Müller (ed.), *(In)vulnerable Domains in Multilingualism*. Amsterdam: Benjamins, 295-330.

- Moravcsik, Edith. 1978. 'Universals of language contact'. In J. Greenberg (ed.), *Universals of Human Language, vol. 1: Method and theory*. Stanford, CA; Stanford University Press, 93-122.
- Montague, Richard. 1969. 'On the nature of certain philosophical entities'. *The Monist* 53: 159-94.
- Morgan, Jerry. 1972. 'Verb agreement as a rule of English'. In P. Peranteau, J. Levi and G. Phares (eds.), *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 278-286.
- Motyliniski, Gustave-Adolphe. 1904. *Le dialecte berbère de R'edamès*. Paris: Ernest Leroux.
- Mufwene, Salikoko. 2007. 'Population movements and contacts in language evolution'. *Journal of Language Contact THEMA* 1: 63-91.
- Müller, Natascha. 1998. 'Transfer in first language acquisition'. *Bilingualism: Language and Cognition* 1.3: 151-171.
- & A. Hulk. 2001. 'Crosslinguistic influence in bilingual language acquisition'. *Bilingualism: Language and Cognition* 4.1: 1-21.
- Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
- Nadkarni, Mangesh. 1975. 'Bilingualism and syntactic change in Konkani'. *Language* 51: 672-683.
- Nagara, Susumu. 1972. *Japanese Pidgin English in Hawaii: A Bilingual Description*. Honolulu: University Press of Hawaii.
- Newman, Paul. 1980. *The Classification of Chadic within Afroasiatic*. Leiden: Leiden University Press.

- Obler, Lorraine. 1975. 'Reflexes of Classical Arabic šay'un 'thing' in the modern dialects: a study in patterns of language change'. PhD dissertation, University of Michigan.
- 1990. 'Reflexes of Classical Arabic šay'un 'thing' in the modern dialects'. In J. Bellamy (ed.), *Studies in Near Eastern Culture and History in Memory of Ernest T. Abdel-Massih*. Ann Arbor: University of Michigan, 132-152.
- Odlin, Terrence. 1989. *Language Transfer: Cross-linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.
- Orel, Vladimir. 2000. *A Concise Historical Grammar of the Albanian Language*. Leiden: Brill.
- Ouali, Hamid. 2003. 'Sentential negation in Berber: a comparative study'. In J. Mugany (ed.), *Linguistic Description: Typology and Representation of African Languages*. Trenton, NJ: Africa World Press, 243-255.
- Forthcoming. 'Negative expressions in Moroccan Arabic and Modern Standard Arabic: NPI's or NCI's?'. In *Perspectives on Arabic Linguistics: Papers from the Annual Symposium on Arabic Linguistics*, vol. 22. Amsterdam: Benjamins.
- Ouhalla, Jamal. 2002. 'The structure and logical form of sentences in Arabic'. In J. Ouhalla & U. Shlonsky (eds.), *Themes in Arabic and Hebrew Syntax*. Dordrecht: Kluwer Academic Publishers, 299-320.
- Owens, Jonathan. 1984. *A Short Reference Grammar of Eastern Libyan Arabic*. Wiesbaden: Otto Harrassowitz.
- 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Helmut Buske Verlag.

- 1988. *The Foundations of Grammar: An Introduction to Medieval Arabic Grammatical Theory*. Amsterdam: Benjamins.
- 2003. 'Arabic dialect history and historical linguistic mythology'. *Journal of the American Oriental Society* 123: 715-740.
- 2006. *A Linguistic History of Arabic*. Oxford: Oxford University Press.
- Palva, Heikki. 1984. 'A general classification for the Arabic dialects spoken in Palestine and TransJordan'. *Studia Orientalia* 55.18: 359-376.
- 2004. 'Negations in the dialect of es-Salt, Jordan'. In C. Versteegh, M. Woidich, M. Haak & R. de Jong (eds.), *Approaches to Arabic Dialects: A Collection of Articles Presented to Manfred Woidich on the Occasion of his Sixtieth Birthday*. Leiden: Brill, 221-236.
- Paradisi, Umberto. 1961. 'Testi berberi di Augila (Cirenaïca)'. *Annali dell'Istituto Orientale di Napoli* 10: 79-91.
- Parry, Mair. 1997. 'Preverbal negation and clitic ordering, with particular reference to a group of north-west Italian dialects'. *Zeitschrift für romanische Philologie* 113: 244-270.
- Parry, Mair. 2009. 'The development of negation in Italo-Romance'. In Breitbarth et al. (in preparation).
- Paspati, Alexandre. 1870. *Études sur les Tchinghianés ou Bohémiens de l'Empire Ottoman*. Osnabrück: Biblio.
- Pat-El, Na'ama. Forthcoming. 'On negation in Phoenician'. In A. Holmstedt & A. Schade (eds.), *The Phoenician Language: Studies in Writing and Linguistics*. Winona Lake, Indiana: Eisenbrauns.

- Patkanoff, K. P. 1907/8. 'Some words on the dialects of the Transcaucasian gypsies', *Journal of the Gypsy Lore Society* (new series) 1: 229-257.
- Penka, Doris. 2007. 'Negative indefinites'. Ph.D. dissertation, Eberhard Karls Universität Tübingen.
- Picard, André. 1957. 'Du prétérit intensif en berbère'. In 'Comité André Basset' (eds.), *Mémorial André Basset (1895-1956)*. Paris: Adrien-Maisonneuve, 107-120.
- Pienemann, Manfred & G. Håkansson. 2007. 'Full transfer vs. developmentally moderated transfer: a reply to Bohnacker'. *Second Language Research* 23.4: 485-493.
- Poletto, Cecilia. 2007. 'On negation splitting and doubling'. Paper presented at the Grand Meeting for the Network for Scandinavian Dialect Syntax, Lake Mývatn, Iceland.
- Price, Glanville. 1993. '*Pas (point)* without *ne* in interrogative clauses'. *Journal of French Language Studies* 3.2: 191-195.
- Progovac, Liliana. 2006. 'The syntax of nonsententials: small clauses and phrases at the root'. In L. Progovac, K. Paesani, E. Casielles & E. Barton (eds.), *The Syntax of Nonsententials: Multidisciplinary Perspectives*. Amsterdam: Benjamins, 33-71.
- Provotelle, Paul. 1911. *Étude sur la Tamazir't ou, Zénatia de Qalaât Es-Sened (Tunisie)*. Paris: Ernest Leroux.
- Pustet, Regina. 2003. *Copulas: Universals in the Categorization of the Lexicon*. Oxford: Oxford University Press.
- Rabhi, Allahoua. 1996. 'De la négation en berbère: les donnés algériennes'. In Chaker & Caubet (1996), 23-34.

- Randall, Janet. 1992. 'The Catapult Hypothesis: an approach to unlearning'. In J. Weissenborn, H. Goodluck, & T. Roeper (eds.), *Theoretical Issues in Language Acquisition: Continuity and Change in Development*. Hillsdale, NJ: Lawrence Erlbaum, 93-138.
- Reichmuth, Stefan. 1983. *Der arabische Dialekt der Šukriyya im Ostsudan*. Hildesheim: Georg Olms Verlag.
- Reinhardt, Carl. 1894. *Ein arabischer Dialekt gesprochen in 'Omān und Zanzibar: nach praktischen Gesichtspunkten für das Seminar für orientalische Sprachen in Berlin*. Berlin: W. Spemann.
- Reintges, Chris. 2004. *Coptic Egyptian (Sahidic Dialect): A Learner's Grammar*. Köln: Rüdiger Köpper Verlag.
- Reymond, Eve & J. Barns. 1973. 'Introduction'. In E. Reymond & J. Barns (eds.), *Four Martyrdoms from the Pierpoint Morgan Coptic Codices*. Oxford: Clarendon Press.
- Robert, Stéphane. 2005. 'The challenge of polygrammaticalization for linguistic theory'. In Z. Frajzyngier, A. Hodges & D. Rood (eds.), *Linguistic Diversity and Language Theories*. Amsterdam: Benjamins, 119-142.
- Roberts, Ian. 1994. 'Universal Grammar and First Language Acquisition'. In N. Ellis (ed.), *Implicit and Explicit Learning of Languages*. New York: Academic Press, 455-476.
- 2001. 'Language change and learnability'. In S. Bertolo (ed.), *Parametric Linguistics and Learnability: a Self Contained Tutorial for Linguists*. Cambridge: Cambridge University Press, 81-125.
- 2007. *Diachronic Syntax*. Oxford: Oxford University Press.

- Rosén, Haim. 1977. *Contemporary Hebrew*. The Hague: Mouton.
- Rosenhouse, Judith. 1984. 'Towards a classification of Bedouin dialects in Israel'.
Bulletin of the School of Oriental and African Studies 47: 508-522.
- 2003/4. 'ilo/'indo/ma'ó used to express possession in colloquial Arabic in Israel'. *Mediterranean Language Review* 15: 129-153.
- Rubin, Aaron. 2005. *Studies in Semitic Grammaticalization*. Winona Lake: Eisenbrauns.
- Ruhlen, Merritt. 1994. *On the Origin of Languages: Studies in Linguistic Taxonomy*. Palo Alto: Stanford University Press.
- Saeed, John. 1999. *Somali*. Amsterdam: Benjamins.
- Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt Brace.
- Schachter, Jacquelyn. 1996. 'Maturation and the issue of Universal Grammar in second language acquisition'. In W. Ritchie & T. Bhatia (eds.), *Handbook of Second Language Acquisition*. New York: Academic Press, 159-194.
- Schmid, Monika. 2002. *First Language Attrition, Use and Maintenance: The Case of German Jews in Anglophone Countries*. Amsterdam: Benjamins.
- Schuchardt, Hugo. 1884. *Slawo-deutsches und Slawo-italienisches*. Graz: Leuschner & Lubensky.
- 1885. *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin.
- Schütze, Carson. 1996. *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.
- Schwartz, Bonnie. 1998. 'The second language instinct'. *Lingua* 106: 133-160.

- & R. Sprouse. 1994. 'Word order and nominative case in non-native language acquisition: a longitudinal study of (L1 Turkish) German interlanguage'. In T. Hoekstra & B. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar: Papers in Honor of Kenneth Wexler from the 1991 GLOW Workshop*. Amsterdam: Benjamins, 317-368.
- 1996. 'L2 cognitive states and the Full Transfer/Full Access model'. *Second Language Research* 12: 40-77.
- Schwegler, Armin. 1988. 'Word-order changes in predicate negation strategies in Romance languages'. *Diachronica* 5.1/2: 21-58.
- Schwenter, Scott. 2006. 'Fine-tuning Jespersen's Cycle'. In B. Birner & G. Ward (eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*. Amsterdam: Benjamins, 327-344.
- Segal, Moses. 2001. *A Grammar of Mishnaic Hebrew*. Oxford: Oxford University Press.
- Seliger, Herbert. 1996. 'Primary language attrition in the context of bilingualism'. In W. Ritchie & T. Bhatia (eds.), *Handbook of Second Language Acquisition*. New York: Academic Press, 605-626.
- Shahin, Kimary. 2000. *Rural Palestinian Arabic (Abu Shusha Dialect)*, Munich: Lincom Europa.
- Sharwood Smith, Michael & P. van Buren. 1991. 'First language attrition and the parameter setting model'. In H. Seliger & R. Vago (eds.), *First Language Attrition*. Cambridge: Cambridge University Press, 17-30.
- Shaw, Thomas. 1738. *Travels, or, observations relating to several parts of Barbary and the Levant*. Oxford: printed at the theatre.

- Silva-Corvalán, Carmen. 1998. 'On borrowing as a mechanism of syntactic change'.
In A. Schwegler, B. Tranel & M. Uribe-Etxebarria (eds.), *Romance Linguistics: Theoretical Perspectives*. Amsterdam: Benjamins, 225-246.
- Simeone-Senelle, Marie-Claude. 1997. 'The Modern South Arabian languages'. In Hetzron (1997a), 378-423.
- Smits, Caroline. 1998. 'Two models for the study of language contact: a psycholinguistic perspective versus a socio-cultural perspective'. In M. Schmidt, J. Austin & D. Stein (eds.), *Historical Linguistics 1997: Selected papers from the 13th International Conference on Historical Linguistics, Düsseldorf, 10-17 August 1997*. Amsterdam: Benjamins, 377-391.
- Soane, E. B. 1913. *Grammar of the Kurmanji or Kurdish Language*. London: Luzac.
- Solihin, Sohirin. 1991. *Copts and Muslims in Egypt: A Study on Harmony and Hostility*. Leicester: The Islamic Foundation.
- Sperber, Dan & D. Wilson. 1995. *Relevance: Communication and Cognition*. Oxford: Blackwell, 2nd edition.
- Spitta-Bey, Wilhelm. 1880. *Grammatik des arabischen vulgärdialectes von Aegypten*. Leipzig: J.C. Hinrichs.
- Stainton, Robert. 2005. 'In defense of non-sentential assertion', in Z. Szabo (ed.), *Semantics Versus Pragmatics*. Oxford: Oxford University Press, 383-457.
- Stassen, Leon. 2008a. 'Predicative possession'. In: Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 117. Available online at <http://wals.info/feature/117>. Accessed on 2009-09-28.

- 2008b. 'Comparative constructions'. In M. Haspelmath, M. Dryer, D. Gil, & B. Comrie (eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 121. Available online at <http://wals.info/feature/121>. Accessed on 2009-09-28.
- Stauble, Anne-Marie. 1984. 'A comparison of a Spanish-English and a Japanese-English second language continuum: negation and verb morphology'. In R. Andersen (Ed.), *Second Languages: A Cross-Linguistic Perspective*. Rowley, Mass.: Newbury House, 323-353.
- Stechow, Arnim von. 1984. 'Comparing semantic theories of comparison'. *Journal of Semantics* 3: 1-77.
- Sutcliffe, Edmund. 1936. *A Grammar of the Maltese Language, with Chrestomathy and Vocabulary*. London: Oxford University Press.
- de Swart, Henriëtte. 2006. 'Marking and interpretation of negation: a bi-directional Optimality Theory approach'. In R. Zanuttini, H. Campos, E. Herburger & P. Portner (eds.), *Crosslinguistic Research in Syntax and Semantics: Negation, Tense and Clausal Architecture*. Washington: Georgetown University Press, 199-218.
- 2008. 'Negative polarity and negative indefinites in an optimization approach: a sketch of features, typology and grammaticalization'. Handout from a talk presented at the first *Cycles of Grammaticalization* seminar, Aston University, June 2008.
- & I. Sag. 2002. 'Negation and negative concord in Romance'. *Linguistics and Philosophy* 25.4: 373-417.

- Taha, Abdulwahid. 1989. *The Muslim Conquest and Settlement of North Africa and Spain*. London: Routledge.
- Talmon, Rafi. 1999. 'The syntactic study of *maf'ūl muṭlaq*: a study in Qur'ānic syntax'. In Y. Suleiman (ed.), *Arabic Grammar and Linguistics*. London: Curzon Press.
- Theil, Rolf. 2006. 'Is Omotic Afro-Asiatic?'. Ms., University of Oslo.
- Thomason, Sarah. 1983. 'Genetic relationship and the case of Ma'a'. *Studies in African Linguistics* 14: 195-231.
- 2007. 'Language contact and deliberate change'. *Journal of Language Contact* *THEMA* 1: 41-61.
- & T. Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Tieken-Boon van Ostade, Ingrid, G. Tottie, & W. van der Wurff (eds.) 1998. *Negation in the History of English*. Berlin: Mouton de Gruyter.
- Timberlake, Alan, 1977. 'Reanalysis and actualization in syntactic change'. In C. Li (ed.), *Mechanisms of Syntactic Change*. Austin: University of Texas Press, 141-177.
- Tonciulescu, Keren. 2007. 'Indefinites and negative polarity items in Modern Hebrew'. *Cahiers linguistiques d'Ottawa* 35: 183-206.
- Trips, Carola. 2002. *From OV to VO in Early Middle English*. Amsterdam: Benjamins.
- Tropper, Josef. 2002. *Altäthiopisch: Grammatik des Ge'ez mit Übungstexten und Glossar*. Münster: Ugarit-Verlag.
- Tsiapera, Mária. 1969. *A Descriptive Analysis of Cypriot Maronite Arabic*. The Hague: Mouton.

- Tsimpli, Ianthi & A. Roussou. 1991. 'Parameter resetting in L2?'. *UCL Working Papers in Linguistics* 3: 149-169.
- Tsimpli, Ianthi, A. Sorace, C. Heycock & F. Filiaci, 2004. 'First language attrition and syntactic subjects: a study of Greek and Italian near-native speakers of English'. *International Journal of Bilingualism* 8.3: 257-277.
- Turner, Lorenzo. 2002. *Africanisms in the Gullah dialect*. Columbia, SC: University of South Carolina Press.
- van der Auwera, Johan. 2009. 'The Jespersen Cycles' In Elly van Gelderen (ed.), *Cyclical Change*. Amsterdam: Benjamins, 35-70.
- In press. 'On the diachrony of negation'. In L. Horn (ed.), *The Expression of Negation*. Berlin: Mouton de Gruyter.
- & A. Neuckermans. 2004. 'Jespersen's Cycle and the interaction of predicate and quantifier negation in Flemish'. In B. Kortmann (ed.), *Typology Meets Dialectology: Dialect Grammar from a Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter, 454-79.
- Verhaar, John. 1995. *Towards a Reference Grammar of Tok Pisin: An Experiment in Corpus Linguistics*. Honolulu: University of Hawai'i Press.
- Volterra, Virginia & T. Taeschner. 1978. 'The acquisition and development of language by bilingual children'. *Journal of Child Language* 5: 311-326.
- Vrolijk, Arnoud. 1998. *Bringing a Laugh to a Scowling Face: A Study and Critical Edition of the 'Nuzhat al-Nufūs wa-Muḍḥik al-'abūs' by 'Alī Ibn Sūdūn al-Bašbuḡāwī*. Leiden: CNWS Publications.
- Wagner, Ewald. 1983. *Harari-Texte in arabischer Schrift*. Wiesbaden: Franz Steiner Verlag.

- 1997. 'Harari', in Hetzron (1997a), 486-508.
- Wagner, Esther-Miriam. 2007. 'A linguistic analysis of Judaeo-Arabic letters from the Cairo Genizah'. PhD dissertation, University of Cambridge.
- Wallage, Phillip. 2005. 'Negation in Early English: parametric variation and grammatical competition'. PhD dissertation, University of York.
- Walsh, Terrence & K. Diller. 1981. 'Neurolinguistic considerations on the optimum age for second language learning'. In K. Diller (ed.), *Individual Difference and Universals in Language Learning Aptitude*. Rowley, Mass.: Newbury House, 3-21.
- Watanabe, Akira. 2004. 'The genesis of negative concord: syntax and morphology of negative doubling'. *Linguistic Inquiry* 35.4: 559-612.
- Watson, Janet. 1993. *A Syntax of Ṣan'āni Arabic*. Wiesbaden: Otto Harrassowitz.
- 2002. *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.
- Watterson, Barbara. 1988. *Coptic Egypt*. Edinburgh: Scottish Academic Press.
- Weerman, Fred. 1993. 'The diachronic consequences of first and second language acquisition: the change from OV to VO'. *Linguistics* 31: 903-931.
- Wehr, Hans. 1979. *A Dictionary of Modern Written Arabic*, vol. 4. J.M. Cowan (ed.). Wiesbaden: Harrassowitz.
- Weinreich, Uriel. 1953. *Languages in Contact: Findings and Problems*. New York: Linguistic Circle of New York. (Reprinted 1986, The Hague: Mouton).
- , W. Labov, & W. Herzog. 1968. 'Empirical foundations for a theory of language change'. In W. Lehmann & Y. Malkiel (eds.), *Directions for Historical Linguistics*. Austin: University of Texas Press, 95-195.

- Wensinck, Arent. 1955. *Concordance et indices de la tradition musulmane*, vol. 3. Leiden: Brill.
- Wexler, Ken. 1998. 'Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage'. *Lingua* 106: 23-79.
- White, Lydia. 1991. 'Adverb placement in second language acquisition: some effects of positive and negative evidence in the classroom'. *Second Language Research* 7.2: 133-161.
- Willis, David. 2006. 'Negation in Middle Welsh'. *Studia Celtica* 40.1: 63-88.
- To appear. 'A Minimalist approach to Jespersen's Cycle in Welsh'. In D. Jonas, J. Whitman & A. Garrett (eds.), *Grammatical Change: Origins, Nature and Outcomes*. Oxford: Oxford University Press.
- Forthcoming. 'Motivating the emergence of new markers of sentential negation: the case of Welsh *ddim*'. *Diachronica*.
- Willmore, John. 1901. *The Spoken Arabic of Egypt*. London.
- Winand, Jean. 1997. 'La négation *bn... iwn3* en Néo-Égyptien'. *Lingua Aegyptia* 5: 223-236.
- Winford, Donald. 2005. 'Contact-induced changes: classification and processes'. *Diachronica* 22.2: 373-427.
- 2007. 'Some issues in the study of language contact'. *Journal of Language Contact THEMA* 1: 22-39.
- Woidich, Manfred. 1968. 'Negation und negative Sätze im Ägyptisch-Arabischen'. Ph.D. dissertation, Ludwig-Maximilians-Universität zu München.
- 2006. *Das Kairenisch-Arabische: Eine Grammatik*. Wiesbaden: Harrassowitz Verlag.

- Wurff, Wim van der. 1998. 'On expletive negation with adversative predicates in the history of English'. In Tieken-Boon van Ostade et al. (1998), 295-327.
- Zanuttini, Raffaella. 1989. 'Two strategies for negation: evidence from Romance'. In J. Powers and K. de Jong (eds.), *Proceedings of ESCOL V*, Ohio State University, 535-546.
- Zeijlstra, Hedde. 2004. 'Sentential negation and negative concord'. PhD dissertation, University of Amsterdam. Utrecht: LOT Publications.
- 2005. 'What the Dutch Jespersen Cycle may reveal about negative concord'. In M. Andronis, E. Debenport, A. Pycha & K. Yoshimura (eds.), *Proceedings of the 38th Meeting of the Chicago Linguistics Society*, vol. 2. Chicago: Chicago Linguistics Society, 143-158.
- Zima, Petr. 2007. 'Why languages *and* contacts?'. *Journal of Language Contact* *THEMA* 1: 101-115.
- Zuckermann, Gil'ad. 2009. 'Hybridity versus revivability: multiple causation, forms and patterns'. *Journal of Language Contact* 2: 40-67.
- Zwarts, Frans. 1995. 'Nonveridical contexts'. *Linguistic Analysis* 25: 286-312.
- 1998. 'Three types of polarity'. In F. Hamm & E. Hinrichs (eds.), *Plurality and Quantification*. Dordrecht: Kluwer Academic Publishers, 177-238.

Appendix: questionnaire on Palestinian Arabic negation

شكراً جزيلاً لمساعدتك في مشروع البحث هذا. من المقصود بهذا الاستبيان أنني أتعلم كيف تعبر أنت شخصياً عن نفسك **بلهجتك** (وليس بلغة العربية الفصحى).

Thank you very much for your help in this research project. The intention of this questionnaire is for me to learn how you personally express yourself in your **own dialect** of spoken Arabic (not in standard written Arabic).

المقدمة

Preliminaries

أ - الاسم:

a) Name:

ب - مكان الولادة:

b) Place of birth:

ج - مكان السكن الحالي:

c) Current place of residence:

د - أماكن أخرى سكنت فيها:

d) Other places you have lived:

هل شكل كل جملة من الجمل التالية عادي أم غريب أم مستحيل بلهجتك؟

Is the **form** of each of the following sentences normal, odd or impossible in your dialect?

معنى الجمل غير مهم.

The meaning of the sentences is not important.

مثلاً:

معنى الجملة "السمك يمشي على أرض" غريب/خطأ بالرغم من أن شكلها عادي. ولاكن شكل الجملة "الزلمة بتحب بلدها" غريب/مستحيل بالرغم من أن معناها عادي.

For example:

The meaning of the sentence "Fish walk on land" is strange or incorrect, but its form is fine. However, the form of the sentence "The man loves (fem.) her country" is strange or impossible, even if its meaning is fine.

الأسئلة:

Questions:

1. ما فش عصير في المطبخ.

1. mā fi-š ʕaṣīr fi-l maṭbax.
NEG there.is-NEG juice in-the kitchen
'There isn't any juice in the kitchen.'

عادي

Normal

غريب

Odd

مستحيل

Impossible

2. فش عصير في المطبخ.

2. fi-š ʕaṣīr fi-l maṭbax.
there.is-NEG juice in-the kitchen
'There isn't any juice in the kitchen.'

Normal	عادي
Odd	غريب
Impossible	مستحيل

3. ما في عصير في المطبخ.

3. mā fī ṣaṣīr fi-l maṭbax.
 NEG there.is juice in-the kitchen
 'There isn't any juice in the kitchen.'

Normal	عادي
Odd	غريب
Impossible	مستحيل

4. (أنا) ما بحبّش الفول.

4. (ana) mā baḥibb-iš il-fūl.
 I NEG like.IMPF.1S-NEG the-fava.beans
 'I don't like fava beans.'

Normal	عادي
Odd	غريب
Impossible	مستحيل

5. (أنا) بحبّش الفول.

5. (ana) baḥibb-iš il-fūl.
 I like.IMPF.1S-NEG the-fava.beans
 'I don't like fava beans.'

Normal

عادي

Odd

غريب

Impossible

مستحيل

6. (أنا) ما بحب الفول.

6. (ana) mā baḥibb il-fūl.
 I NEG like.IMP.F.1S the-fava.beans
 'I don't like fava beans.'

Normal

عادي

Odd

غريب

Impossible

مستحيل

7. (أنا) ما أكلتش الفول.

7. (ana) mā akalt-iš il-fūl.
 I NEG eat.PRF.1S-NEG the-fava.beans
 'I didn't eat the fava beans.'

Normal

عادي

Odd

غريب

Impossible

مستحيل

8. (أنا) أكلتش الفول.

8. (ana) akalt-iš il-fūl.
 I eat.PRF.1S-NEG the-fava.beans
 'I didn't eat the fava beans.'

Normal

عادي

Odd

غريب

Impossible

مستحيل

9. (إنت) أكلتش الفول؟ (سؤال)

9. (inte) akalt-iš il-fūl? [suʔāl]
 you eat.PRF.2MS-Q(.NEG) the-fava.beans [question]
 'Didn't you eat the fava beans?' or 'Did you eat the fava beans?'

Normal

عادي

Odd

غريب

Impossible

مستحيل

10. ما مسحناش الطاولة.

10. mā mesaḥnā-š iṭ-ṭāwila.
 NEG wipe.PRF.1P-NEG the-table
 'We didn't wipe the table.'

Normal

عادي

Odd

غريب

Impossible

مستحيل

11. مسحناش الطاولة.

11. mesaḥnā-š iṭ-ṭāwila.
 wipe.PRF.1P-NEG the-table
 'We didn't wipe the table.'

Normal	عادي
Odd	غريب
Impossible	مستحيل

12. اذا شفتش أحمد سلم عليه.

12. iza	šuft-š	aḥmad	sallim	ʔalē.
if	see.PRF.2MS-š	Ahmad	greet.IMPER.2MS	upon.him
'If you see Ahmad, greet him.'				

Normal	عادي
Odd	غريب
Impossible	مستحيل

13. اشتريت شي عدس وشي زيت.

13. ištarēt	šī	ʔadas	u-šī	zēt.
buy.PRF.1S	some	lentils	and-some	oil
'I bought some lentils and some oil.'				

Normal	عادي
Odd	غريب
Impossible	مستحيل

14. اشتريت اشى عدس واشى زيت.

14. iṣṭarēt iṣi ḥadas w-iṣi zēt.
 buy.PRF.1S some lentils and-some oil
 'I bought some lentils and some oil.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

15. ما بديش مصاري.

15. mā biddī-š maṣāri.
 NEG want.1S-NEG money
 'I don't want money.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

16. بديش مصاري.

16. biddī-š maṣāri.
 want.1S-NEG money
 'I don't want money.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

17. ما معيش مصاري.

17. mā maʿī-š maṣāri.
NEG have.1S-NEG money
'I don't have money.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

18. معيش مصاري.

18. maʿī-š maṣāri.
have.1S-NEG money
'I don't have money.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

19. ما عنديش مصاري.

19. mā ʿandī-š maṣāri.
NEG have.1S-NEG money
'I don't have money.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

20. عنديش مصاري.

20. ʕandī-š mašāri.
have.1S-NEG money
'I don't have money.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

21. عمري ما أكلت لحم الخنزير.

21. ʕumr-i mā akalt laḥm il-xinzīr.
age-my NEG eat.PRF.1S meat the-pig
'I have never eaten pork.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

22. عمري ما أكلتش لحم الخنزير.

22. ʕumr-i mā akalt-iš laḥm il-xinzīr.
age-my NEG eat.PRF.1S-NEG meat the-pig
'I have never eaten pork.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

23. والله ما بحب الفول.

23. wallāhi mā baḥibb il-fūl.
by.God NEG like.IMPF.1S the-fava.beans
'By God, I really don't like fava beans!'

Normal

Odd

Impossible

عادي

غريب

مستحيل

24. والله ما بحبش الفول.

24. wallāhi mā baḥibb-iš il-fūl.
by.God NEG like.IMPF.1S-NEG the-fava.beans
'By God, I really don't like fava beans!'

Normal

Odd

Impossible

عادي

غريب

مستحيل

25. والله بحبش الفول.

25. wallāhi baḥibb-iš il-fūl.
by.God like.IMPF.1S-NEG the-fava.beans
'By God, I really don't like fava beans!'

Normal

Odd

Impossible

عادي

غريب

مستحيل

26. شفت أحمد امبارح شي؟

26. šuft aḥmad imbāriḥ ši.
see.PRF.2MS Ahmad yesterday šī
'Did you see Ahmad yesterday (at all)?'

Normal

Odd

Impossible

عادي

غريب

مستحيل

27. اذا شفت أحمد بكرة شي سلم عليه.

27. iza šuft aḥmad bukra ši sallim ʔalē.
if see.PRF.2MS Ahmad tomorrow šī greet.IMPER.2MS upon.him
'If you see Ahmad tomorrow (at all), greet him.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

28. ما عمريش رحت الصين.

28. mā ʔumr-ī-š ruḥt iṣ-šīn.
NEG age-my-NEG go.PRF.1S the-China
'I have never been to China.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

29. عمریش رحٲ الصین.

29. ʕumr-ī-š ruḥt iṣ-ṣīn.
age-my-NEG go.PRF.1S the-China
'I have never been to China.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

30. ما شفناش إشي

30. mā šufnā-š iši.
NEG see.PRF.1P-NEG thing
'We didn't see a thing.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

31. ما شفناش ولا إشي.

31. mā šufnā-š wala iši.
NEG see.PRF.1P-NEG not.even thing
'We didn't see a thing.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

32. ما شفنا إشي

32. mā šufnā iši.
 NEG see.PRF.1P thing
 'We didn't see a thing.'

Normal

Odd

Impossible

عادي

غريب

مستحيل

33. ما شفنا ولا إشي.

33. mā šufnā wala iši.
 NEG see.PRF.1P not.even thing
 'We didn't see a thing.'

Normal

Odd

Impossible

عادي

غريب

مستحيل