Journal of Maltese Studies 30

On Maltese Syntax

Journal of Maltese Studies 30

On Maltese Syntax

EDITED BY

MICHAEL SPAGNOL
RAY FABRI

Department of Maltese
University of Malta
in collaboration with
midsea BOOKS

Published by the Department of Maltese
Faculty of Arts, University of Malta

in collaboration with

Midsea Books Ltd
Carmelites Street, Sta Venera, Malta
www.midseabooks.com

All enquiries should be addressed to:
The Editorial Board, Department of Maltese
Faculty of Arts, University of Malta, Msida, MALTA

First published 2023

Produced by Midsea Books Ltd.

ISBN:978-99932-7-957-0

CONTENTS

PREFACE

This volume of the *Journal of Maltese Studies* aims at presenting the reader with a selection of studies focusing on selected issues in the morpho-syntax of Maltese, with contributions by linguists who are currently active in research.

The aims of this volume are threefold: a) to showcase current research on Maltese syntax, both for a general and for a specialised audience, b) to provide an update of current, state-of-the-art descriptions and analyses of a selected set of topics in syntax in order to stimulate further research within these areas, especially among young scholars of Maltese linguistics, and c) to provide a general introduction to the study of the specific areas chosen while placing the study within a larger picture, and setting the stage for further studies in other related areas.

Maltese has a long and intriguing history going back to its Arabic roots in the 11th century. Through the centuries up to the present, it has gone through phases of intense contact with non-Arabic languages, mainly Sicilian, Italian and English, that have sculpted its unique character and moulded it into a language that has achieved the status of national language and, together with English, official language of the Republic of Malta. Maltese is spoken by a large majority of the Maltese population on a daily basis and boasts a rich literature and a diverse media landscape.[1]

1 Information about the most recent National Statistics Office survey (2021)

Since at least the 17th century, many scholars and travellers have shown an interest in describing the grammar and vocabulary of Maltese, the earliest being the *Thesaurus Polyglottus,* a multilingual dictionary by the German linguist and historian Hieronymus Megiser published in 1603 which features 121 items from Maltese (Cowan 1964). In recent times, there has been a surge of interest by scholars and young researchers interested in exploring various features covering the core areas of language, i.e., phonology, morphology, syntax, semantics, and pragmatics, as well as applied areas of language use, such as language acquisition, language teaching, language contact, and, most importantly in the modern age, the development of digital language resources and tools. Unfortunately, however, Maltese still lags behind in the availability of basic resources such as electronic lexicons, and spell and grammar checkers. This crucial lack of resources is certainly not due to a lack of interest or expertise but to a lack of consistent and targeted financial and human resources that are dedicated to long-term national projects specifically focused on developing such crucial tools.

Maltese has a Maghrebi Arabic stratum, a Romance superstratum (Sicilian, Italian) and an English adstratum. According to Brincat (2000, p. 24), "we cannot decide whether the substrate of Maltese should be Punic, Latin or Greek, for the simple reason that in the Maltese language there is no substrate" (our translation). As a result of intensive language contact, Arabic Maltese has undergone a process of relexification, first through contact with Sicilian, and later Florentine Italian, followed by English, which is currently the main source of borrowing.

The morphosyntax of Maltese retains a strong Arabic character, although this is difficult, if not impossible, to quantify, with lexical items borrowed from Sicilian, Italian, and English

commissioned by the National Council for the Maltese Language is available on http://www.kunsilltalmalti.gov.mt/news-details?nwid=223&ctid=19&ctref=kollaborazzjoni

either being formally integrated into the Arabic root system to some extent or other, such as the 2nd form (binyan, declension) verb *pejjep* 'to smoke' from Italian *pipa* 'pipe', or retaining the stem and inflecting it through Arabic based affixes, such as *niskorja* 'I score (a goal)' and *tiskorja* 'you score (a goal)' from English 'score (a goal)' (see Mifsud 1995 for an in-depth study of loan verbs). Contact from two very different language families, Semitic and Indo-European, has thus resulted in an intriguing mixed morphological system displaying both root and stem bases through integration and innovation.

In its syntax, Maltese also displays typical Arabic features, mostly reminiscent of Maghrebi varieties, in particular Tunisian Arabic, in their expression, such as the construct state, e.g. *xagħar it-tifel* 'the boy's hair, lit. hair the boy', and nominal sentences, e.g. *Ħija tabib* 'My brother is a doctor, lit. My brother doctor'. However, it also shows innovations, such as the analytical passive, e.g. *Il-ktieb ġie ppubblikat* 'The book was published', as opposed to the synthetic passive, e.g. *Il-ktieb inkiteb minn awtur żagħżugħ* 'The book was written by a young author' (see in particular Lukas & Čéplö 2020 for a discussion of contact-induced changes in Maltese, and Ebert 2000 for a discussion of TMA forms in Arabic and Maltese).

The six articles in the present volume set out to provide a description and analysis of a number of salient (morpho-) syntactic constructions and phenomena of Modern Maltese from a synchronic perspective. The first two articles are pitched at the clause level. Albert Borg's contribution deals with nominal complement clauses and in various positions (subject, object, in apposition, as complement to adjective, preposition and adverb) in both declaratives and interrogatives as well as with various types of adverbial clauses (time, manner, conditional, etc.), exploring in detail the rich variety displayed and the differences and similarities between them. Maris Camilleri focuses on the relative clause, specifically restrictive, non-restrictive, and free

relative clauses, focusing on the strategies adopted (resumptive pronouns, gaps, etc.) and also highlighting interesting differences between Standard and Dialect. Together, these two contributions make up an in-depth study of the clause in Maltese.

These are followed by two contributions that focus on two specific syntagmatic relations, i.e. the relation between elements within the string that makes up a phrase or a clause. Ray Fabri goes into the details of grammatical agreement in Maltese within two main syntactic domains, namely, within the noun phrase (demonstrative - noun, attributive adjective - noun, etc.) and outside of the noun phrase (verb - subject, verb - object, predicative adjective - subject, etc.), also touching upon cases of long-distance pronoun - antecedent agreement. The contribution ends with a discussion of examples of notional agreement, which, at least on the surface, involves two elements that do not agree formally within a domain in which they would normally agree formally, thus forcing a specific reading based on the semantics of the elements involved, therefore, 'semantic agreement'.

Slavomír Čéplö looks at constituent order, in terms of S(ubject), V(erb) and O(bject) within declarative sentences, critically reviewing previous analyses of Maltese. Based on an analysis of corpus data, he concludes that Maltese 'looks…more like a strict SVO language like English', with the only exception being SV in existential clauses. Given these two contributions, it would be interesting to explore in more detail the relation between agreement and constituent order, assuming, of course, that such a relation exists.

Finally, the remaining two contributions deal with very specific phenomena. Christopher Lucas describes in great detail various constructions involving negation in Maltese, both at the sentence/clause (main and subordinate) level and at the subsentential (constituent) level. He explores the relation between negation and indefinite pronouns and ends by discussing the occurrence of the suffix *-x*, which is typically used to negate verbs, in non-negative contexts, for which he offers an interesting, plausible explanation.

Thomas Stolz, Nataliya Levkovtch, and Maike Vorholt investigate the occurrences of the three spatial propositions which express Place or Goal (in, at, inside, within), namely, *fi*, *ġo* and *ġewwa* as well as 'zero-marking', when they take place names as complements, trying to work out on the basis of what criteria (e.g., familiarity or complexity of the place name) a particular preposition is chosen rather than another, or nothing. The article sets a solid basis for more analyses of the prepositional system of Maltese, and raises a number of questions hopefully to be taken up in future research by scholars interested in Maltese, in particular, and in syntax and syntactic theory in general.

Taken together, these articles cover important areas of the Maltese (morpho-)syntactic landscape. Clearly, there is still a great deal of research that needs to be done; however, we hope that this volume can serve as an incentive for more scholars to explore the grammar of Maltese, and come up with descriptions and theoretical explanations of the observed phenomena. The volume is intended as the first in a series of publications by the University of Malta, covering not only topics in Maltese syntax but also in the other areas of linguistic analysis, both core, ie., phonetics, phonology, morphology, semantics, pragmatics, and lexicography, and applied areas, such as language acquisition, sociolinguistics, and historical linguistics.

We thank all the authors for their invaluable contributions (and patience) to this volume, and also our two research assistants, Raffaello Bezzina and Michela Vella, for their help in various stages of production of this volume. Of course, any errors remain our responsibility.

Ray Fabri & Michael Spagnol
Malta, 24th January, 2023

References

Brincat, Joseph M (2011) *Maltese and Other Languages: a linguistic history of Malta*. Malta: Midsea

Cowan, William (1964) An Early Maltese Word List. *Journal of Maltese Studies 2*, 217-225.

Ebert, Karen (2000) Aspect in Maltese, in Dahl, Östen (ed.) *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter

Lukas, Christopher & Čéplö, Slavomír (2020) Maltese. In Lukas, Christopher & Manfredi, Stefano (eds) *Arabic and Contact-Induced Change*. Berlin: Language Science Press

Mifsud, Manwel (1995) *Loan Verbs in Maltese: A Descriptive and Comparative Study*. New York: Brill

NOMINAL AND ADVERBIAL CLAUSES IN MALTESE[1]

Albert Borg

Abstract

The article considers two types of subordinate sentences: nominal (complement) clauses and adverbial clauses. It touches upon the difference between a nominal clause with an explicit verb and a clause lacking an explicit verb (a predicate/copular sentence, also known as a nominal (non-verbal) sentence), before going on to explore declarative and question-word interrogative nominal clauses, yes-no interrogative and alternative interrogative nominal clauses as well as headless relative clauses. Adverbial clauses are next considered, starting with a look at simple and compound adverbial conjunctions. Various types of clauses are discussed: adverbial clauses of time, manner, purpose, result and cause/reason. The study goes on to look at conditional adverbial clauses, concessive clauses, alternative concessive adverbial clauses, comparative and equative adverbial clauses. The article concludes with a look at adverbial clauses which are rendered non-finite through nominalisation.

1 I would like to thank my colleagues Marie Azzopardi Alexander and Ray Fabri for their helpful comments on various aspects of this paper. There is a lot in the present work which draws upon the earlier volume by Borg and Azzopardi Alexander (1997). Of course I am responsible for the views expressed here.

Dan l-artiklu jagħti ħarsa lejn żewġ tipi ta' sentenzi subordinati: sentenzi subordinati nominali kumplamentari u sentenzi subordinati avverbjali. Jibda billi jittratta fuq fuq id-differenza bejn sentenza subordinata li tinkludi verb u sentenza subordinata bla verb (sentenza kopulari jew predikattiva, magħrufa wkoll bħala sentenza nominali). Imbagħad jifli sentenzi subordinati nominali dikjarattivi, sentenzi interrogattivi b'espressjoni interrogattiva, mistoqsijiet li jitolbu t-tweġiba iva jew le, mistoqsijiet alternattivi u sentenzi subordinati aġġettivali mingħajr ras. Jeżamina wkoll sentenzi subordinati avverbjali billi qabelxejn jagħti ħarsa lejn konġunzjonijiet avverbjali sempliċi u komposti. Jistudja tipi differenti ta' sentenzi subordinati avverbjali: ta' żmien, ta' manjiera, ta' skop, ta' riżultat u ta' kawża jew raġuni. Imbagħad ikompli billi jqis sentenzi subordinati avverbjali kondizzjonali, sentenzi konċessivi, sentenzi konċessivi alternattivi, sentenzi komparattivi u sentenzi avverbjali ekwattivi. Jagħlaq b'diskussjoni fuq sentenzi subordinati avverbjali li jsiru mhux finiti permezz tan-nominalizzazzjoni.

1. Nominal clauses

To date not much is known, within the field of Maltese syntax, about subordinate clauses. While Camilleri (this volume) treats relative clauses, the present study deals, if only in a preliminary way, with nominal (complement) clauses and adverbial clauses. A subordinate sentence (or clause) is called "nominal" when it serves the function of a noun phrase within a matrix sentence. In Maltese, such (subordinate) nominal clauses are usually introduced with the subordinating conjunction *li* which also introduces (subordinate) adjectival (i.e relative) clauses. This subordinate status is also borne out by the intonation contour over the clause in question: the native speaker intuits that it is not a (syntactically) complete structure (cf the description of the intonation contour of the nominal clause in example 2 below).

In what follows, I distinguish a subordinate nominal clause from nominal sentences (cf Borg 1987-88 for different types of nominal sentences in Maltese) which, however, can also be subordinate. A nominal sentence, such as example (1), has a subject (*dan*) and a predicate (*inaċċettabbli*) in which there is no verb and yet its intonation contour marks it as complete.

(1) *Dan* *in-aċċettabbli*.
this.SGM un-acceptable
'This is unacceptable.'

Completeness in declarative Maltese sentences is often marked by an intonation contour that shows the final stressed syllable - *inaċċet**tab**bli* - to have a falling intonation. This is often followed by a rise in the following unstressed syllable - *inaċċettabb**li*** (Azzopardi-Alexander, personal communication).

For reasons not entirely as yet specified, a small range of expressions may appear between the subject and the predicate, depending on the type of nominal sentence involved. One such expression is a form of the independent pronoun used with copular function. In the case of (1), the form is the third person singular masculine pronoun *huwa* or *hu*, thus, *Dan huwa/hu in-aċċettabbli.*

A nominal clause, as indicated, is typically introduced by the subordinating conjunction *li* "that":

(2) *Li* *t-idħaq* *f'-din* *is-sitwazzjoni,*
that 2-laugh.IPFV.SG in-this.SGF DEF-situation
in-aċċettabbli.
un-acceptable
'That you should laugh in this situation is unacceptable.'

In this example it occupies the position of subject of the (main) sentence in which it occurs, substituting for the demonstrative pronoun *dan* (the subject of the sentence in 1). The incomplete and, therefore, subordinate status of the nominal clause is marked by a rising intonation on the last stressed syllable – *sitwazz**jo**ni*

- that continues till the end of the word. This gives the listener the expectation of a continuation which will be marked by a falling intonation on the final unstressed syllable of that element (in the main clause) that completes it and, possibly a rise on the unstressed syllable as in (1) (Azzopardi-Alexander, personal communication).

It is to be noted also that a subordinate nominal clause may itself be a nominal, as opposed to a verbal, sentence. Compare example (3) with (4):

(3) *Oħt-u* *(hija)* *s-segretarja.*
sister-his (COP.SGF) DEF-secretary
'His sister is the secretary.'

(4) *Li* *oħt-u* *(hija)* *s-segretarja* *(huwa)*
that sister-his (COP.SGF) DEF-secretary (COP.3SGM)
in-aċċettabbli.
un-acceptable
'It is unacceptable that his sister is the secretary.'

1.1 Declarative nominal clauses

A nominal clause can also occur as the object of a (verbal) sentence:

(5) *Id-difiża* *argument-at* *li* *l-imputat*
DEF-defence argue.PFV-3SGF that DEF-accused
tfixkel *f-it-tapit.*
trip up.PFV.3SGM in-DEF-carpet
'The defence argued that the accused tripped up on the carpet.'

Let us go back to nominal clauses in subject position: we have already seen a nominal clause as subject of a nominal sentence (examples 1 and 3): however when it comes to a nominal clause as subject of a verbal sentence, it seems that the subordinating conjunction *li* has to be expanded to *Il-fatt li...* "The fact that..." (but cf discussion of left dislocation in example (13) below).

(6) *Il-fatt* *li* *għed-t-l-u,* *qawwie-l-u*
DEF-fact that say.PFV-1SG-OM-him strengthen.PFV.3SGM-OM-him
qalb-u.
heart-his
'The fact that I told him served to encourage him.'

The nominal clause *li għedtlu* is an argument of *Il-fatt* so that the subject of (6) is *Il-fatt li għedtlu*.

Although one can also have *Il-fatt li* introducing the subject nominal clause of the nominal sentence in (2), it is not clear why its use should be felt to be more required in the case of the subject of a verbal sentence like (6).

There are a number of verbs which typically take a (often contiguous) subordinate nominal clause for object: *qal* "he said", *ħolom* "he dreamed", *stqarr* "he declared", *emmen* "he believed", *argumenta* "he argued", etc. The nominalisation corresponding to each of these verbs can also take a following contiguous subordinate nominal clause. One could argue that the transitivity of the verb is preserved in the nominalisation: this would make the nominal clause an argument (object) of the nominalised verb. Examples (7) and (8) illustrate an instance of this:

(7) *Stqarr-et* *li* *ma* *t-af* *xejn.*
state.PFV-3SGF that NEG 3F-know.IPFV.SG nothing
'She declared that she did not know anything.'

(8) *L-istqarrija* *li* *ma* *t-af* *xejn*
DEF-statement that NEG 3F-know.IPFV.SG nothing
ma *emmin-ha* *ħadd.*
NEG believe.PFV-3SGM nobody
'Nobody believed her statement that she did not know anything.'

Now there are other nouns which are not verbal nouns and they can also take a following subordinate nominal clause, as in:

(9)	*Ix-xniegħa*	*li*	*ġġarraf*	*il-ħajt*
	DEF-rumour	that	was.destroyed.PFV.3SGM	DEF-wall
	ħassb-et	*lil*	*kulħadd.*	
	worry.PFV-3SGF	OM	all	
	'Everybody was worried at the rumour that the wall had collapsed.'			

It would seem that in a case such as example (9), the following nominal clause can also be analysed as being an argument of the subject noun phrase (as in the case also of examples (6) and (8)).

Since the subordinating conjunction *li* introduces both nominal and adjectival clauses, the question arises as to how one can distinguish between the two types in this context, following a noun phrase.

One solution is to have recourse to the presence of implicit or explicit co-referentiality between the noun phrase and the following clause, necessarily present in the case of an adjectival clause. Thus, in (10) the subject of the subordinate clause is co-referential with the subject of the main clause, clearly making the subordinate clause an adjectival one.

(10)	*Ix-xniegħa*	*li*	*ħassb-et*	*lil*	*kulħadd*	*hija*
	DEF-rumour	that	worry.PFV-3SGF	OM	everyone	COP.3SGF
	falz-a.					
	false-SGF					
	'The rumour which had everybody worried is false.'					

On the other hand no element of the subordinate clause in (9) *li ġġarraf il-ħajt* is co-referential with the subject of the main sentence: in this case we clearly have a nominal, not an adjectival, clause.

Still one can have a subordinate clause with one of its arguments co-referential with an argument in the main clause, while still being a nominal, not an adjectival, clause, as in

(11)	*Ulied-u*	*qal-u*	*li*	*se*	*j-itilq-u.*
	children-his	say.PFV-3PL	that	FUT	3-leave.IPFV-PL
	'His children said that they would be leaving.'				

Here the subject of the main clause and that of the subordinate clause are co-referential, but the subordinate clause is clearly a nominal one (the subordinate clause does not occur in a noun phrase and does not have a noun to modify). It therefore seems that there is also a semantic principle involved: the nominal clause "spells out" the intended content of the noun or verb it is complement to or object of. In (9) the nominal clause gives the "content" of the noun: what the rumour is about, and in (11) the nominal clause spells out the intended content of the verb of saying. In contrast, the adjectival clause in (10) simply characterises the rumour mentioned in a particular way.

A nominal clause can, in addition, be in apposition to an object noun phrase as in example (12):

(12) *Semgħu-ha* *din* *ix-xniegħa,* *li*
hear.PFV.3PL-3SGF this.SGF DEF-rumour that
arrest-aw *li-s-suspettat.*
arrest.PFV-3PL OM-DEF-suspect
'They heard this rumour, namely, that they had arrested the suspect.'

We can also have left dislocation of the nominal clause in apposition with what one might regard as a resumptive subject expression, the singular masculine demonstrative pronoun *dan* "this":

(13) *Li* *ġġarraf* *il-ħajt,* *dan*
that destroyed.PFV.3SGM DEF-wall this.SGM
ħasseb *lil* *kulħadd.*
worry.PFV.3SG OM everyone
'That the wall had collapsed was what had everybody worried.'
(Literally, 'That the wall had collapsed, that worried everybody')

However this type of construction does not come across as very natural.

A nominal clause may also occur as complement to an adjective (14) or as complement to a preposition (15):

(14) *It-tfal* *imdejq-in* *li* *tilef* *it-tim*
DEF-children sad-PL that lose.PFV.3M DEF-team
tagħ-hom.
of-3PL
'The children are sad because their team lost.'

(15) *Wara* *li* *kiel* *kemm* *felaħ,*
after that eat.PFV.3SG how.much can.PFV.3SGM
qabad *u* *telaq.*
get.up.PFV.3SGM and leave.PFV.3SGM
'After stuffing himself, he just got up and left.'

A nominal clause can also occur as complement to an adverb:

(16) *Aktarx* *(li)* *se* *j-siefer* *dalwaqt.*
probably (that) FUT 3M-travel.IPFV.SG soon
'He's probably soon going abroad.'

In the case of example (16) the subordinator *li* is optional. In general, one may note that the order of the nominal clause relative to the expression it is complement to, seems restricted to the position following such an expression. Deviations from this order are quite marked and require certain structural adjustments as seen in the case of left dislocation in example (13).

1.2 Question-word question nominal clauses

An interrogative question-word nominal clause may occur as the subject of the sentence as in examples (17) and (18):

(17) *Fejn* *is-siefer* *j-iddepend-i* *mil-l-gost-i*
where 2-travel.IPFV.SG 3M-depend.IPFV-SG from-DEF-taste-PL
tiegħek.
of-2SG
'The choice of places you want to visit abroad depends on your tastes.'

(18) *Kif* *in-ġib-u* *ruħ-na* *j-influwenz-a*
how 1-bring.IPFV-PL soul-1PL 3M-influence.IPFV-SG
r-relazzjoni *tagħ-na* *ma-l-oħr-ajn.*
DEF-relation of-1PL with-DEF-other-PL
'The way we behave has a bearing on our relationship with others.'

It may also occur as the object of the sentence as in examples (19, 20) and as the predicate of a nominal sentence (21, 22):

(19) *Ma* *n-ista-x* *n-ifhem* *x'*
NEG 1-can.IPFV.SG-NEG 1-understand.IPFV.SG what
ġab-ek *hawn.*
bring.PFV.3SGM-2SG here
'I cannot understand what brought you here.'

(20) *Ħi-ja* *saqsie-ni* *fejn* *marr-et/*
brother-1SG ask.PFV.3SGM-1SG where go.PFV-3SGF/
kif *waqaj-t/* *kemm* *infaq-t/* *meta*
how fall.PFV-1SG/ how.much spend.PFV-1SG/ when
se *j-iġ-i.*
FUT 3M-come.IPFV-SG
'My brother asked me where she went to/how I fell/how much I spent/when will he be coming.'

(21) *Il-kwistjoni* *koll-ha* *hi* *x'* *se* *n-agħmel*
DEF-question all-SGF COP.3SGF what FUT 1-do.IPFV.SG
bi-h.
with-3SGM
'The whole point is what should I do with him.'

Such a clause can also occur in apposition to a subject noun phrase (22) or in apposition to an object noun phrase (23).

(22) *Il-punt* *kruċjali,* *x'* *se* *n-agħml-u* *bi-h,*
DEF-point crucial what FUT 1-do.IPFV-PL with-3SGM
għad-u *qed* *j-inkweta-na.*
still-3SGM PROG 3-worry.IPFV.SG-1PL
'The crucial point, what we are to do with him, is still bothering us.'

(23) *M'* *għand-ek-x* *idea* *kemm* *n-ieħu*
NEG at-2SG-NEG idea how.much 1-take.IPFV.SG
paċenzja *bi-h.*
patience with-3SGM
'You have no idea how patient I am with him.'

Question-word nominal clauses may occur also as a complement to an adjective, as in examples (24) and (25):

(24) *M'* *inie-x* *ċert* *x'* *se* *j-agħmel.*
NEG 1SG-NEG certain what FUT 3M-do.IPFV.SG
'I'm not sure what he intends on doing.'

(25) *Int-om* *żgur-i* *kemm* *se* *j-iġ-u*
2-PL certain-PL how.much FUT 3-come.IPFV-PL
mistedn-in?
guest-PL
'Do you know for sure how many guests are coming?'

They may also occur as a complement to a preposition as in examples (26) and (27):

(26) *Ma* *ftehem-nie-x* *fuq* *min* *se* *n-istiedn-u.*
NEG agree.PFV-1PL-NEG on who FUT 1-invite.IPFV-PL
'We have not agreed about whom to invite.'

(27) *Id-dubju* *dwar* *x'* *għamel* *b-il-flus*
DEF-doubt about what do.PFV.3SGM with-DEF-money
se *j-ibqa'* *magħ-na.*
FUT 3M-remain.IPFV.SG with-1PL
'Our doubt about what he did with the money will linger on.'

1.3 Yes-no interrogative nominal clauses

Yes-no interrogative nominal clauses can occur as the object of a sentence:

(28)	*Saqsie-ni*	*(jekk)*	*Pietru*	*kellim-x*	*lil*
	ask.PFV.3SGM-1SG	(if)	Peter	talk.PFV.3SGM-NEG	OM
	ħu-h.				
	brother-3SGM				
	'He asked me whether Peter had spoken to his brother.'				

Note in this nominal clause the optional presence of the complementizer *jekk* 'if, whether' together with the non-negative suffix on the verb[2] (cf Lucas: this volume). We can also have an object yes-no clause which is a nominal (rather than a verbal) sentence:

(29)	*Saqsie-ni*		*(jekk)*	*l-arloġġ*	*hu-x*
	ask.PFV.3SGM-1SG		(if)	DEF-clock	COP.3SGM-NEG
	(qiegħed)	*fuq*	*il-mejda.*		
	(located.3SGM)	on	DEF-table		
	'He asked me whether the clock was on the table.'				

Once again the complementiser *jekk* is optional and there is also the obligatory presence of the element *hux* (with copular function and non-negative suffix *-x*). The locative present participle is optional. In some dialects the present participle *qigħedx* (with non-negative suffix *–x*) could occur instead of *hux*.

A yes-no clause can also occur as the subject of the sentence (30a)[3] and in apposition to a subject noun phrase (30b):

(30a)	*Jekk*	*t-ista-x*	*t-iġi*	*jew*	*le*
	if	2-can.IPFV.SG-NEG	2-come.IPFV.SG	or	no
	ma	*j-interessa-ni-x.*			
	NEG	3M-concern.IPFV.SG-1SG-NEG			
	'Whether you can come or not is of no concern to me.'				

2 For convenience, this non-negative suffix is still glossed as 'neg' in the illustrative sentences

3 Example kindly provided by Ray Fabri

(30b)	*Il-kwistjoni*	*jekk*	*Pietru*	*kellim-x*		*lil*
	DEF-question	if	Peter	talk.PFV.3SGM-NEG		OM
	ħu-h		*qed*	*t-inkweta-na.*		
	brother-3SGM		PROG	3F-worry.IPFV.SG-1PL		
	'The question whether Peter did talk to his brother is troubling us.'					

Note that in this case it is barely possible to omit the complementizer *jekk*.

Such a nominal clause can also occur as the predicate of a nominal sentence:

(31)	*Jekk*	*Pietru*	*kellim-x*	*lil*	*ħu-h*	*hija*
	if	Peter	talk.PFV.3SGM-NEG	OM	brother-3SGM	COP.SGF
	possibbiltà		*reali.*			
	possibility		real			
	'Whether Peter talked to his brother is a distinct possibility.'					

In this sentence, the subject noun phrase is *possibbiltà reali* following, rather than preceding, the predicate, and the copular expression is singular feminine, agreeing with it. Also, as in the case of example (30), the complementizer *jekk* is obligatory.

A yes-no nominal clause can also occur in apposition to an object noun phrase (*d-dubju*), although a case could also be made for treating the clause as its complement:

(32)	*Semmie-l-i*		*d-dubju*	*(jekk)*	*Pietru*
	mention.PFV.3SGM-TO-1SG		DEF-doubt	(if)	Peter
	kien-x	*kellem*	*lil*	*ħu-h.*	
	be.PFV.3SGM-NEG	talk.pfv.3SGM	OM	brother-3SGM	
	'He mentioned the doubt whether Peter had talked to his brother.'				

We can also have such clauses occurring as a complement to a predicative adjective (33) or as a complement to a preposition (34).

(33)	*M'*	*aħnie-x*	*żgur-i*	*(jekk)*	*għamilnie-x*	*sew.*
	NEG	1.PL-NEG	certain-PL	(if)	do.PFV.1PL-NEG	right
	'We are not sure whether we acted correctly.'					

(34) *Ma* *qbilnie-x* *dwar* *jekk* *għandnie-x*
NEG agree.PFV.1PL-NEG about if should.PFV.1PL-NEG
n-itilq-u.
1-leave.IPFV-PL
'We did not agree about whether we should leave.'

Once again note that in example (34) the complementizer *jekk* is obligatory.

1.4 Alternative interrogative nominal clauses

Alternative interrogative nominal clauses can occur as the object of a sentence:

(35) *Saqsie-ni* *(jekk)* *n-ixtieq-x* *n-itlaq*
ask.PFV.3SGM-1SG (if) 1-wish.IPFV.SG-NEG 1-leave.IPFV.SG
jew *n-ibqa'.*
or 1-remain.IPFV.SG
'He asked me whether I wanted to go or leave.'

As in the case of yes-no interrogative nominal clauses, note the optionality of the complementizer *jekk* 'if, whether', and the non-negative suffix on the verb.

Such clauses can also be a nominal rather than a verbal sentence as in example (36) (cf example 29):

(36) *Saqsie-ni* *(jekk)* *l-arloġġ* *hu-x*
ask.PFV.3SGM-1SG (if) DEF-clock COP.3SGM-NEG
(qiegħed) *fuq* *il-mejda* *jew* *fuq*
(located.3SGM) on DEF-table or on
il-gradenza.
DEF-chest.of.drawers
'He asked me whether the clock was on the table or on the chest of drawers.'

Note once more the obligatory copular element in (36) with the non-negative suffix.

An alternative nominal clause can occur as subject of the sentence (37a)[4], and in apposition to a subject noun phrase (37b):

(37a) *Jekk* *mar* *jew* *ma* *mar-x*
if go.PFV.3SGM or NEG go.PFV.3SGM-NEG
ma *j-interessani-x.*
NEG 3M-concern.IPFV.SG-1SG-NEG
'Whether he went or not is no concern of mine.'

(37b) *Il-mistoqsija* *jekk* *n-ixtieq-x* *n-itlaq*
DEF-question if 1-wish.IPFV.SG-NEG 1-leave.IPFV.SG
jew *n-ibqa',* *ikkonfond-iet-u.*
or 1-remain.IPFV.SG confuse.PFV-3SGF-3SGM
'The question whether I want to leave or stay confused him.'

An alternative nominal clause can also occur as the predicate of a nominal sentence:

(38) *Il-mistoqsija* *hi* *jekk* *n-ixtieq-x*
DEF-question COP.3SGF if 1-wish.IPFV.SG-NEG
n-itlaq *jew* *n-ibqa'.*
1-leave.IPFV.SG or 1-remain.IPFV.SG
'The question is whether I would like to leave or stay.'

Sentence (39) exemplifies an alternative nominal clause in apposition to an object noun phrase:

(39) *Semmie-l-i* *d-dubju* *(jekk)* *Pietru*
mention.PFV.3SGM-TO-1SG DEF-doubt (if) Peter
kien-x *kellem* *lil* *ħu-h* *jew* *le.*
be.PFV.3SGM-NEG talk.PFV.3SGM OM brother-3SGM or no
'He mentioned the doubt whether Peter had talked to his brother or not.'

In example (40) the alternative nominal clause occurs as a complement to an adjective and in example (41) it occurs as complement to a preposition.

4 Example kindly provided by Ray Fabri

(40) *M'* *aħnie-x* *żgur-i* *(jekk)* *kellnie-x*
NEG 1.PL-NEG certain-PL (if) have.PFV.1PL-NEG
tort *jew* *raġun.*
guilt or right
'We are not sure whether we were right or wrong.'

(41) *Ma* *qbil-nie-x* *dwar* *jekk* *għand-nie-x*
NEG agree.PFV-1PL-NEG about if at-1PL-NEG
n-itilq-u *jew* *n-oqogħd-u.*
1-leave.IPFV-PL or 1-stay.IPFV-PL
'We did not agree about whether we should leave or stay.'

1.5 Headless relative clauses

There is a class of clauses which seem to imply an antecedent noun phrase, which however is not realised, making them "headless". An alternative term would be "nominal relative clause" (Quirk et al, 1985: 1056). Compare examples (42) and (43) in which the clause *Min jikkommetti reat* can be substituted for the (subject) noun phrase *Il-kriminal.*

(42) *Il-kriminal* *i-rid* *i-ħallas* *ta'*
DEF-criminal 3M-want.IPFV.SG. 3M-pay.IPFV.SG of
għemil-u.
deed-3SGM
'The criminal should pay for his deeds.'

(43) *Min* *j-ikkommett-i* *reat* *i-rid*
who 3M-commit.IPFV-SG crime 3M-want.IPFV.SG
i-ħallas *ta'* *għemil-u.*
3M-pay.IPFV.SG of action-3SGM
'Whoever commits a crime has to pay for his deeds.'

A more complex example is the following:

(44) *T-af* *li* *x'* *qal-l-i* *koll-u*
2-know.IPFV.SG that what say.PFV.3SGM-TO-1SG all-3SGM
nsej-t-u?
forget.PFV-1SG-3SGM
'Do you know that I have forgotten all that he told me?'

This complex sentence clearly involves one main clause *Taf* and two subordinate clauses *li kollu nsejtu* and *x'qalli*. The clause *x'qalli* is the object of the verb *nsejtu* in the second subordinate clause, which in turn is the object of the main verb *taf*. The clause *x'qalli* can be easily substituted by a noun phrase such as *il-fatt* "the fact". At the same time an antecedent noun phrase such as *dak (li qalli)* "that (which he told me)" seems to be implied.

2. Adverbial clauses

Adverbial clauses serve the function of an adverb or adverbial phrase. Characteristically they are introduced by adverbial conjunctions and typically express circumstantial information about the proposition encoded by the main sentence. In example (45) the adverbial clause of purpose, introduced by the adverbial conjunction *biex* "in order to", gives the motivation for the situation encoded by the main sentence:

(45) *Oħt-u* *marr-et* *f-il-kju* *biex* *ma*
sister-3SGM go.PFV-3SGF in-DEF-queue to NEG
t-ibqa-x *l-aħħar*.
3F-remain.IPFV.SG-NEG DEF-last
'His sister queued up so as not to remain last.'

Adverbial clauses may follow the main sentence as in (45) or precede it as in (46), depending on the communicative effect the speaker wants to convey:

(46) *Biex* *ma* *t-ibqa-x* *l-aħħar,* *oħt-u*
to NEG 3F-remain.IPFV.SG-NEG DEF-last sister-3SGM
marr-et *fi-l-kju*.
go.PFV-3SGF in-DEF-queue
'So as not to remain last, his sister queued up.'

Such clauses may even occur between the subject of the main sentence and its predicate:

(47)	*Oħt-u,*	*biex*	*ma*	*t-ibqa-x*	*l-aħħar;*
	sister-3SGM	to	NEG	3F-remain.IPFV.SG-NEG	DEF-last
	marr-et	*fi-l-kju*.			
	go.PFV-3SGF	in-DEF-queue			
	'His sister, not to remain last, queued up.'				

2.1 Adverbial conjunctions

Adverbial conjunctions which figure in the following examples may be simple, e.g. *biex* "in order to" in examples (45-47), *meta* "when", *kif* "how", and others. Adverbial conjunctions may also be compound: *billi* "because, since", *malli* "as soon as" and *talli* "because of" combine a preposition with the element *li* and are conventionally written as one word (but see also the point made below in section **2.12 Non-finite adverbial clauses**, following example 94). Other compound expressions are *minħabba li* "because", *waqt li* "during, while", *wara li* "after that", *tant li* "so much so that", *fil-ħin li* "at the moment that" and *kull darba li* "every time that". Although in spontaneous speech there does not seem to be a pause between this *li* and the preceding expression, nonetheless in these cases, *li* is written as a separate word.

Still other conjunctions combine various elements: *skont kif*, "how" (literally, "according-to how"), *hekk kif* "just as" (literally, "so how"), *bil-mod kif* "according to" (literally, "in-the-way how"), *biex b'hekk* "so that" (literally, "in-order with so"), *kull x'ħin* "every time that" (literally "every what time"), *daqs kemm* "as much as" (literally "as-much how-much").

2.2 Adverbial clauses of time

Different conjunctions give different time specifications. Thus *meta* "when" and *x'ħin* "when" (literally, "what time") indicate

a moment in time expressed in the adverbial clause generally contemporary with the time expressed in the main clause:

(48) *Sakkar* *il-bieb* *meta* *t-oħroġ.*
lock.PFV.2SG DEF-door when 2-go.out.IPFV.SG
'Lock the door when you go out.'

(49) *Sellm-it-l-u* *x'ħin* *ra-t-u.*
greet.PFV-3SGF-to-3SGM what-time see.PFV-3SGF-3SGM
'She greeted him when she saw him.'

The conjunctions *fil-ħin li* "when, at the time that, at the very moment that", *malli* "as soon as" and *kif/hekk kif* (in this context) "as soon as, exactly when" also express contemporaneity, but with the added notion of immediacy:

(50) *Inħb-ej-na* *kif* *ra-j-nie-h* *ġej.*
hide.PFV-PL-1PL how see.PFV-PL-1PL-3SGM coming.SGM
'We hid as soon as we saw him approaching.'

(51) *Ħarb-et* *malli* *ra-t-na.*
escape.PFV-3SGF as.soon.as see.PFV-3SGF-1PL
'She escaped as soon as she saw us.'

Kull x'ħin/meta, *kull darb-a li* "every time that" express a moment of time in the subordinate clause distributively commensurate with the corresponding moment in the main clause:

(52) *N-olqot* *minkb-i* *kull* *darb-a* *li*
1-hit.IPFV.SG elbow-1SG every time-SGF that
n-idħol *f'di-l-karozza.*
1-enter.IPFV.SG in-this.SGF-DEF-car
'I graze my elbow every time I get into this car.'

The conjunction *qabel ma* expresses a time posterior to that expressed in the main clause, while *wara li* expresses a time anterior to that in the main clause:

(53) *Leħq-et* *telq-et* *qabel* *ma* *sellm-il-ha*.
reach.PFV-3SGF leave.PFV-3SGF before NEG greet.PFV.3SGM-to-3SGF
'She had already left before he managed to salute her.'

(54) *Iffirma* *l-kuntratt* *wara* *li* *qra-h*
sign.PFV.3SGM DEF-contract after that read.PFV.3SGM-3SGM
koll-u.
all-3SGM
'He signed the contract after reading it in full.'

In contrast, in the case of the conjunction *waqt li*, the time expressed in the main clause is embedded within that expressed by the adverbial clause:

(55) *Ħabbat* *il-bieb* *waqt* *li*
knock.PFV.3SGM DEF-door during that
kien *qed* *j-iekol*.
be.PFV.3SGM PROG 3M-eat.IPFV.SG
'There was a knock at the door while he was eating.'

In the case of the conjunctions *sakemm/sa ma*, the time expressed in the adverbial clause marks the boundary of the expiry of the time expressed in the main clause:

(56) *Se* *n-żomm-l-ok* *post* *sakemm* *t-asal*.
FUT 1-keep.IPFV.SG-TO-2SG place until 2-arrive.IPFV.SG
'I'll keep you a place till you arrive.'

Conversely, with the conjunction *(sa) (minn) mindu*, the time expressed in the adverbial clause expresses the starting point of the time expressed by the main clause:

(57) *Ħad-et* *ħsieb-u* *sa* *minn* *mindu*
take.PFV-3SGF thought-3SGM to from since
kien *tarbija*.
be.PFV.3SGM baby
'She took care of him from the time he had been a baby.'

2.3 Adverbial clauses of manner

An adverbial clause of manner is typically introduced by the conjunction *kif* "how":

(58)	*Ħawwel*	*is-siġr-a*	*kif*	*urie-h*
	plant.PFV.3SGM	DEF-tree-SGF	as	show.PFV.3SGM-3SGM
	il-ġardinar.			
	DEF-gardener			
	'He planted the tree just as the gardener showed him how to.'			

Skont kif "according to how"and *bil-mod kif* "in the way that" are two possible elaborations of the same conjunction with equivalent meaning. The element *hekk* "so"in the compound conjunction *hekk kif* "exactly as", however, introduces the idea of a more precise manner in the adverbial clause, (*hekk* "so" stresses the congruence of what is expressed in the main clause with the manner indicated in the subordinate clause):

(59)	*Waħħal*	*il-bolla*	*hekk*	*kif*	*qall-u*
	stick.PFV.3SGM	DEF-stamp	so	as	say.PFV.3SGM-3SGM
	j-agħmel	*ta-l-posta*.			
	3M-do.IPFV.SG	of-DEF-post			
	'He stuck the stamp just as the clerk at the post office told him how to.'				

Note also that in this case the adverbial clause can only follow the main clause.

2.4 Adverbial clauses of purpose

The conjunction *biex* "in order to", as seen in example (45) above, introduces a purpose clause which gives the motivation for the situation encoded by the main sentence. There is also the compound conjunction *biex b'hekk* "so that in this way" which, as in the case of *hekk kif* (example 59), introduces an idea of preciseness and deliberateness:

(60) *Qata'* *l-biljett* *biex* *b'* *hekk* *i-kun*
cut.PFV.3SGM DEF-ticket to with so 3M-be.IPFV.SG
j-ista' *j-itlaq* *għal* *Ruma* *l-lejl-a* *stess*.
3M-can.IPFV.SG 3M-leave.IPFV.SG for Rome DEF-night-F EMPHASIS
'He has bought his ticket so he can leave for Rome this very evening.'

Note that in this case too, the adverbial clause can only follow the main clause.

2.5 Adverbial clauses of result

There are clear instances when the same conjunction *biex* introduces an adverbial clause indicating result:

(61) *Għej-a* *wisq* *biex* *j-erġa'*
get.tired.PFV-3SGM too.much to 3M-repeat.IPFV.3SG
j-ikkompet-i.
3M-compete.IPFV-SG
'He got too tired to be able to compete again.'

For result clauses one typically finds the composite expression *tant...li* "so much...that", each element of which introduces a clause as in example (62):

(62) *It-tfal* *tant* *mexx-ej-nie-hom*
DEF-children so.much cause.to.walk.PFV-1PL-3PL
li *f-l-aħħar* *għej-ew*.
that in-DEF-last get.tired.PFV-3PL
'We made the children walk so much that at last they got tired.'

Going by the semantics, the result clause is clearly the one introduced by the conjunction *li*. The rest of the complex sentence *It-tfal tant mexxejniehom* is a transitive sentence with a topicalised object *It-tfal* (cf Borg, Albert and Azzopardi Alexander 2009) . This should be the main clause, except that if we go by the intonation over it, we get the contour associated with incompleteness. For the purposes of this article, I am considering it as the main clause.

Another conjunction expressing result is *tant kemm* "so much", but in this case the element *li* is omitted before the subordinate clause:

(63)	*Tant*	*kemm*	*mexx-ej-nie-hom,*
	so.much	how.much	cause.to.walk.PFV-1PL-3PL
	it-tfal	*f-l-aħħar*	*għej-ew*.
	DEF-children	in-DEF-last	get.tired.PFV-3PL
	'We made them walk so much, the children at last got tired.'		

Whereas the relative order of the main clause followed by the adverbial clause is fixed in (62), in (63) it is not.

It is also interesting to note that in some cases of coordination, the second conjoined sentence can express result:

(64)	*Waqa'*	*u*	*kiser*	*sieq-u*.
	fall.PFV.3SGM	and	break.PFV.3SGM	foot-3SGM
	'He fell and broke a leg.'			

The specific temporal succession of the two situations identified by the two conjoined sentences is fixed by their relative order, signifying the latter as the result of the former.

2.6 Adverbial clauses of cause (or reason)

The conjunction *għax* "because" typically introduces an adverbial clause of reason:

(65)	*Ħareġ*	*kmieni*	*għax*	*kell-u*	*ħafna*	*x'*
	leave.PFV.3SGM	early	because	have.PFV-3SGM	much	what
	j-agħmel.					
	3M-do.IPFV.SG					
	'He left early because he was very busy.'					

Other conjunctions introducing clauses of reason are *billi* "since", *ġaladarba* "once that/given that", *minħabba li* "because of".

Yet another conjunction relevant in this context is *sakemm*, literally "until" but in this context, "as long as":

(66)	*Sakemm*	*ma*	*t-imlie-x*	*il-formula,*	*ma*
	until	NEG	2-fill.IPFV.SG-NEG	DEF-form	NEG
	t-iħ-u-x		*is-sussidju.*		
	2-take.IPFV-SG-NEG		DEF-subsidy		
	'As long as you do not complete the form, you will not receive the subsidy.'				

The subordinate clause can be interpreted as referring to a particular achievement to take place at a given moment, but it can also be interpreted as the (critical) reason or cause making possible the realisation of the situation identified by the main clause.

2.7 Conditional adverbial clauses

Traditionally a distinction is made between the expression of "real" (realis) and "unreal" (irrealis) conditions, and we can also find this formally marked in Maltese to a certain extent. Typically the conjunction introducing a real condition is *jekk* "if", as in the following example:

(67)	*Jekk*	*t-agħmel*	*ix-xita,*	*il-ħamrija*
	if	3F-do.IPFV.SG	DEF-rain	DEF-soil
	t-irtab.			
	3F-become.soft.IPFV.SG			
	'If it rains, the soil will become soft.'			

The conjunction *kieku* "if" typically introduces an unreal condition:

(68)	*Kieku*	*għaml-et*	*ix-xita,*	*il-ħamrija*
	if	do.PFV-3SGF	DEF-rain	DEF-soil
	kien-et	*t-irtab.*		
	be.PFV-3SGF	3F-become.soft.IPFV.SG		
	'Had it rained, the soil would have become soft.'			

While for some speakers it is possible to have *kieku* for *jekk* in example (67), *jekk* instead of *kieku* in example (68) is not possible.

The occurrence of *kieku* to introduce a real condition sounds even more plausible if *kieku*, in addition, also introduces the main clause, thereby possibly underlining the conditionality involved:

(69) *Kieku* *t-agħmel* *ix-xita,* *kieku* *l-ħamrija*
if 3F-do.IPFV.SG DEF-rain if(then) DEF-soil
t-irtab.
3F-become.soft.IPFV.SG
'If only it had to rain, the soil would become soft.'

The conditionality of the subordinate clause can be somewhat emphasised through the use of the conjunction *dment li* "as long as" instead:

(70) *Dment* *li* *t-agħmel* *ix-xita,* *il-ħamrija*
while that 3F-do.IPFV.SG DEF-rain DEF-soil
t-irtab.
3F-become.soft.IPFV.SG
'As long as it rains, the soil will become soft.'

A condition can be further emphasised and highlighted through the use of *ukoll* "also" or *anki* "even" preceding the conjunction *jekk*:

(71) *Ukoll/anki* *jekk* *omm* *t-insa* *t-tarbija*
also/even if mother 3F-forget.IPFV.SG DEF-baby
ta' *ġuf-ha,* *jiena* *ma* *n-insie-k-x.*
of womb-3SGF 1SG NEG 1-forget.IPFV.SG-2SG-NEG
'Even if a mother were to forget the child of her womb, I will not forget you.'

And a further grade of emphasis can be achieved by adding the qualifier *xorta* in the main clause:

(72) *Ukoll/anki* *jekk* *omm* *t-insa* *t-tarbija* *ta'*
also/even if mother 3F-forget.IPFV.SG DEF-baby of
ġuf-ha, *jiena* *xorta* *ma* *n-insie-k-x.*
womb-3SGF 1SG same NEG 1-forget.IPFV.SG-2SG-NEG
'Even if a mother were to forget the child of her womb, even then I will not forget you.'

An equivalent effect can be had inserting *ukoll* instead of *xorta* in the main clause, but this usage sounds somewhat archaic:

(73) *Ukoll/anki* *jekk* *omm* *t-insa* *t-tarbija* *ta'*
also/even if mother 3F-forget.IPFV.SG DEF-baby of
ġuf-ha, *ukoll* *jiena* *ma* *n-insie-k-x.*
womb-3SGF also 1SG NEG 1-forget.IPFV.SG-2SG-NEG
'Even if a mother were to forget the child of her womb, just the same, I will not forget you.'

Note that *ukoll* can only be inserted in the main clause if it also introduces the conditional clause. The 'unreality' of the condition can be reinforced by having *li* precede *kieku* (compare 74 with 68):

(74) *Li* *kieku* *għaml-et* *ix-xita,* *il-ħamrija*
that if do.PFV-3SGM DEF-rain DEF-soil
kien-et *t-irtab.*
be.PFV-3SGF 3F-become.soft.IPFV.SG
'If only it had rained, the soil would have become soft.'

It is to be noted that there are certain restrictions on the choice of tense in both the main and the subordinate clause. Thus the real conditions (67) and (69-73) have an imperfect verb, expressing a time subsequent to the expression of the condition. Correspondingly, the main clause also has an imperfect verb, expressing a time following that of the condition. The real condition can also have future time reference:

(75) *Jekk* *se* *t-issuppervja,* *mhux* *se* *n-kellm-ek.*
if FUT 2-sulk.IPFV.SG NEG FUT 1-speak.IPFV.SG-2SG
'If you are going to sulk, I will not talk to you.'

However, the unreal conditions in (68) and (74) have a perfect verb with past time reference and the main clause also has past time reference subsequent to that expressed by the condition.

2.8 Concessive adverbial clauses

The conditional clauses we have been considering, in relation to their main clause, follow the schema: *'if x, then y'*. Concessive clauses exemplify the schema *'x, but (not) y'*, that is, there is an opposition between *x* and *y* which can also be expressed by the negation of *y*. Typically such clauses are introduced by *minkejja li*, *għalkemm* or *allavolja*:

(76) *Minkejja* *li* *għaml-et* *ix-xita,* *il-ħamrija*
although that do.PFV-3SGF DEF-rain DEF-soil
ma *rtab-it-x.*
NEG become.soft.PFV-3SGF-NEG
'Although it rained, the soil did not become soft.'

Note that the verb in the main clause is negative: the force of the negation in the main clause can be amplified through the use of *xorta* preceding the main verb:

(77) *Minkejja* *li* *għaml-et* *ix-xita,* *il-ħamrija*
although that do.PFV-3SGF DEF-rain DEF-soil
xorta *ma* *rtab-it-x.*
same NEG become.soft.PFV-3SGF-NEG
'Although it rained, the soil still did not soften.'

Time reference with such clauses is not restricted to the past: the subordinate clause can express a condition obtaining generally as in the following example:

(78) *Minkejja* *li* *t-aħdem* *ħafna,* *xorta* *ma*
although that 3F-work.IPFV.SG much same NEG
t-faddal-x *biżżejjed*.
3F-save.up.IPFV.SG-NEG enough
'Although she works a lot, all the same, she doesn't manage to save up enough.'

It can even have future time reference:

(79) *Minkejja* *li* *se* *t-itlaq,* *xorta* *mhux* *se*
although that FUT 2-leave.IPFV.SG same NEG FUT
t-eħles *minn-i*.
2-get.rid.of.IPFV.SG from-1SG
'Although you are leaving, all the same, you will not be getting rid of me.'

As pointed out earlier, one can also have a concessive clause without negation on the main verb:

(80) *Minkejja* *li* *mard-et,* *xorta* *baqa'*
although that get.sick.PFV-3SGF same remain.PFV.3SGM
j-ħobb-ha.
3M-love.IPFV.SG-3SGF
'Although she became ill, he still went on loving her.'

2.9 Alternative concessive clauses

The expression *kemm jekk* or *sew jekk* introduces alternatives within this type of concessive clause:

(81) *Kemm* *jekk* *t-iġ-i* *waħd-ek,* *(u)*
how.much if 2-come.IPFV-SG alone-2SG (and)
kemm *jekk* *t-iġ-i* *ma'* *xi* *ħaddieħor,*
how.much if 2-come.IPFV-SG with some one.else
il-bieb *dejjem* *miftuħ*.
DEF-door always open
'Whether you come on your own or whether you come in somebody else's company, you are always welcome.'

Note the optional occurrence in (81) of the conjunction *u*.

The expression *xorta* can be used in the main clause to highlight the 'concession":

(82)	*Kemm*	*jekk*	*t-iġ-i*		*waħd-ek,*	*(u)*
	how.much	if	2-come.IPFV-SG		alone-2SG	(and)
	kemm	*jekk*	*t-iġ-i*		*ma'*	*xi*
	how.much	if	2-come.IPFV-SG		with	some
	ħaddieħor,	*il-bieb*	*xorta*	*dejjem*	*miftuħ.*	
	one.else	DEF-door	same	always	open	

'Whether you come on your own or whether you come in somebody else's company, just the same, you are always welcome.'

Alternative concessive clauses can have a positive main verb as in (81) and (82), but they can also occur with a negative main verb, as in the following example:

(83)	*Kemm*	*jekk*	*t-ixrob*	*u*	*kemm*	*jekk*
	how.much	if	2-drink.IPVF.SG	and	how.much	if
	t-iekol,	*ma*	*t-ista-x*		*t-esaġera.*	
	2-eat.IPFV.SG	NEG	2-be.able.IPFV.SG-NEG		2-exagerate.IPFV.SG	

'Whether you drink or whether you eat, you cannot overdo it.'

2.10 Comparative adverbial clauses

A comparative adverbial clause can be introduced by one of a small range of adverbial conjunctions: *iktar* or *iżjed*, both meaning "more", and *inqas* "less", followed by the expression *ma* (homonymous with the negative expression *ma*) which can be characterised as a type of relative conjunction. The main clause is also introduced with a corresponding form, as in the following example, so that the main clause, in addition to the subordinate one, contains a comparative expression (correlative comparative):

(84)	*Iktar*	*ma*	*kiel-u,*	*iktar*	*ħxien-u.*
	more	that	eat.PFV-3PL	more	become.fat.PFV-3PL

'The more they ate, the fatter they grew.'

Although semantically both clauses have a comparative element ("eating more', "growing fatter", cf the discussion of example 62), the intonation contour associated with incompleteness is clearly present in the case of the first clause which cannot stand on its own. However - as far as the intonation goes - it is possible to imagine the second clause standing on its own. Accordingly, for present purposes, this is treated as the main clause.

Different combinations of these forms are possible: *iktar ma ... iktar..., iktar ma ... inqas ..., iżjed ma ... iżjed ..., iżjed ma ... inqas ..., inqas ma ... inqas ..., inqas ma ... iktar/iżjed ...* . In the case of some of these combinations it is also possible to have the element *ma* following the conjunction within the main clause:

(85)	*Inqas*	*ma*	*tkellm-et,*		*iżjed*	*(ma)*
	less	that	speak.PFV-3SGF		more	(that)
	nkedd-u.					
	get.annoyed.PFV-3PL					
	'The less she spoke, the more they were annoyed.'					

In general, the preferred order is for the adverbial clause to precede the main clause but it is also possible for the main clause to occur before the subordinate one (with an appropriate intonation contour). However there is a further comparative construction in which deciding which is the main clause is also not such a straightforward matter although on different grounds, as in the following example:

(86)	*Iktar*	*milli*	*ma*	*t-għid*	*xejn,*	*iktar*	*li*
	more	than	NEG	2-say.IPFV.SG	nothing	more	that
	t-uża		*l-prudenza.*				
	2-use.PFV.SG		DEF-prudence				
	'Rather than not saying anything, it's more a question of being prudent.'						

The expression *iktar* is now followed by *milli* and the form *ma* preceding the verb *tgħid* is part of the negative construction. Given the intonation contour over the second clause, it is possible

to imagine it standing on its own. However structural elements are missing here, so one has to posit a case of ellipsis, making the element *iktar li tuża l-prudenza* part of the main clause, as in:

(87) *Iktar* *milli* *ma* *t-għid* *xejn,* *iktar*
more than NEG 2-say.IPFV.SG nothing more
(hija *kwistjoni)* *li* *t-uża* *l-prudenza*.
(COP.3SGF question) that 2-use.PFV.SG DEF-prudence
'Rather than not saying anything, it's more a question of being prudent.'

The main clause would be a nominal sentence whose predicate is *hija kwistjoni* together with an elided subject, so that *li tuża l-prudenza* would be a complement noun clause to the (elided) predicate noun *kwistjoni*.

2.11 Equative adverbial clauses

The compound conjunction *daqskemm* introduces equative clauses:

(88) *Għadd-ew* *mil-l-eżamijiet* *kollha* *daqskemm*
pass.PFV-3PL from-DEF-exams all as.much.as
studj-aw.
study.PFV-3PL
'They studied so much, they passed all their exams.'

Main and subordinate equative clauses may be positive or negative, or they may also be both positive or both negative.

2.12 Non-finite adverbial clauses

Some of the adverbial clauses we have been examining can be made non-finite through nominalisation. Compare the time clause *malli ħarġet il-vara* indicating a particular moment in time in the following example with its nominalisation *mal-ħruġ tal-vara*:

(89) *Bde-w* *j-ispara-w* *il-murtali* *malli*
begin.PFV-3PL 3-let.off.IPFV-PL DEF-fireworks as.soon.as
ħarġ-et *il-vara/* *ma-l-ħruġ*
come.out.PFV-3SGF DEF-statue/ with-DEF-coming.out
ta-l-vara.
of-DEF-statue
'They started letting off fireworks the moment the statue was brought out ('came out')/at the appearance ('coming out') of the statue'.

Note that the conjunction *malli* does not figure in the nominalisation which is now the object of the preposition *ma'* "with" (see also the point made below, following example 94).

The clause *waqt li l-maġistrat kien qed jaqra s-sentenza* indicating a certain duration of time in the following example can also be nominalised to *waqt il-qari tas-sentenza (mill-maġistrat)*.

(90) *L-akkużat* *deher* *kalm* *waqt* *li* *l-maġistrat*
DEF-accused appear.PFV.3SGM calm during that DEF-magistrate
kien *qed* *j-aqra* *s-sentenza/* *waqt*
be.PFV.3SGM PROG 3M-read.IPFV.SG DEF-sentence/ during
il-qari *ta-s-sentenza* *(mil-l-maġistrat)*.
DEF-reading of-DEF-sentence (from-DEF-magistrate)
"The accused seemed calm while the magistrate was reading out the judgement/during the reading out of the judgement (by the magistrate)".

In this case note that the conjunction *waqt li* is reduced to the preposition *waqt* which takes the corresponding nominalisation as its object.

Other possible nominalisations in the case of adverbial clauses of time are the following:

(91) *Żamm-ew-l-ha* *post* *sa* *ma* *wasl-et/*
keep.PFV-3PL-to-3SGF place till that arrive.PFV-3SGF/
sa-l-wasla *tagħ-ha*.
till-DEF-arrival of-3SGF
'They kept a place for her till she arrived/until her arrival.'

The nominalisation *l-wasla tagħha,* object of the preposition *sa,* corresponds to the adverbial clause *sa ma waslet*.

(92) *Ħad-et* *ħsieb-u* *sa* *minn* *mindu* *twieled/*
take.PFV-3SGF thought-3SGM till from since be.born.PFV.3SGM
sa *minn* *twelid-u*.
till from birth-3SGM
'She cared for him from the moment he was born/from (the time of his) birth.'

Here the nominalisation *twelidu*, object of the compound preposition *sa minn* corresponds to the adverbial clause *sa minn mindu twieled*.

The following is an example of the nominalisation of an adverbial clause of manner:

(93) *Dejjem* *mexa* *skont* *kif* *rabb-ew-h*
always walk.PFV.3SGM according how bring.up.PFV-3PL-3SGM
missier-u *u* *omm-u/* *skont* *it-trobbija*
father-3SGM and mother-3SGM/ according DEF-bringing.up
ta' *missier-u* *u* *omm-u*.
of father-3SGM and mother-3SGM
'He always behaved in accordance with how he was brought up by his father and mother/in accordance with his upbringing by his father and mother.'

Note that the nominalised version loses the conjunction *kif*.

In the following example involving the nominalisation of a clause of reason with the compound conjunction *minħabba li*, note the loss of *li*, as in the case of examples (89) and (90):

(94) *Wasal-na* *tard* *minħabba* *li* *ttardj-a*
arrive.PFV-1PL late because.of that be.late.PFV-3SGM
l-ajruplan/ *minħabba* *l-ittardjar* *ta-l-ajruplan*.
DEF-plane/ because.of DEF-lateness of-DEF-plane
'We arrived late because the flight was delayed/because of the delay of the flight.'

In all the cases of nominalisations of adverbial clauses we have seen so far, the conjunction involved was compound, and the process of nominalisation involved the loss of an element in

the conjunction, *li* in (45), (46) and (50), *ma* in (47), *mindu* in (48) and *kif* in (49). It seems these elements are closely bound with the verbal (and finite) status of the clause they introduce. This would explain why they do not figure once the clause is nominalised. Furthermore, in the case of compound conjunctions such as *malli*, *talli*, *filli* it would seem that there is still a synchronic awareness of their composition, since the element *li*, as just seen, is omitted in the nominalised version of the clause.

The situation is different when it comes to the nominalisation of a conditional clause:

(95) *Jekk* *t-irbaħ* *din* *il-battalja/*
if 2-win.IPFV.SG this.SGF DEF-battle/
b-ir-rebħa *ta'* *din* *il-battalja,*
with-DEF-victory of this.SGF DEF-battle
t-ikkonsolida *l-pożizzjoni* *tiegħ-ek.*
2-consolidate.IPFV.SG DEF-position of-2SG
'If you win this battle/by winning this battle, you will consolidate your position.'

The adverbial clause expresses a 'real' condition introduced by the conjunction *jekk* which is simply omitted in the nominalised version. The nominalisation involves a prepositional phrase with the preposition *bi*.

We can also have the nominalisation of an 'unreal' condition introduced by the conjunction *kieku*:

(96) *Kieku* *rbaħ-t* *din* *il-battalja/*
if win.PFV-2SG this.SGF DEF-battle/
b-ir-rebħa *ta'* *din* *il-battalja,* *kon-t*
with-DEF-victory of this DEF-battle be.PFV-2SG
t-ikkonsolida *l-pożizzjoni* *tiegħ-ek.*
2-consolidate.IPFV.SG DEF-position of-2SG
'Had you won this battle/by winning this battle, you would have consolidated your position.'

Note that the nominalisation in (96) via a prepositional phrase with *bi* is identical to that in the 'real' condition in (95) and the conjunction *kieku* is simply left out. The 'unreality' of the

condition is brought out through the use of the (remote) past tense in the main clause, whereas in the case of the 'real' condition in (95), the verb in both the subordinate and the main clause is in the imperfect.

3. Conclusion

The study of these two types of subordinate clauses is still in its infancy. One could perhaps regard the clauses studied here as 'canonical' ones, in the sense that most resemble more or less 'well-formed' sentences produced consciously for the purpose of illustration. However much still needs to be done to account for the transition from actual utterances to (abstract) sentence structures.

Abbreviations

1, 2, 3	first, second, third person	M	masculine
COP	copula	NEG	negative
DEF	definite article	OM	object marker
F	feminine	PFV	perfective
FUT	future	PL	plural
GEN	genitive	PROG	progressive
IPFV	imperfective	SG	singular

References

Borg, Albert (1987-88) To be or not to be a copula in Maltese. *Journal of Maltese Studies* 17–18, 54–71.

Borg, Albert & Azzopardi-Alexander, Marie (1997) *Maltese. Lingua descriptive grammar.* London & New York: Routledge.

Borg, Albert & Azzopardi-Alexander, Marie (2009) Topicalisation in Maltese. In: Comrie, Bernard, Fabri, Ray, Hume, Elizabeth, Mifsud, Manwel, Stolz, Thomas

& Vanhove, Martine (eds.), *Introducing Maltese linguistics*. Amsterdam: John Benjamins, 71–81.

Camilleri, Maris (2023) On relative clauses in Maltese. *Journal of Maltese Studies*. This volume.

Lucas, Christopher (2023) Negation in Maltese. *Journal of Maltese Studies*. This volume.

Quirk, Randolph, Greenbaum, Sidney, Leech, Geoffrey & Svartvik, Jan (1985) *A comprehensive grammar of the English language*. London: Longman.

ON RELATIVE CLAUSES IN MALTESE

Maris Camilleri

Abstract

This work synthesises the literature that makes reference to the relative clause in Maltese, and shows that the relative clause is not a homogeneous structure in the language. Three types of clauses are discussed: restrictive relatives, non-restrictive relatives and free relatives. These come along with their individual constraints both on the antecedent (when available) and on the different strategies they employ. A clear divide between the Standard and dialectal Maltese is shown to exist in the employment of the pronominal strategy, at least in non-free relative clause structures. The discussion also reveals how the availability of complementiser-headed free relatives in Maltese constitutes a rare typological occurrence. This overview of our current knowledge on Maltese relative clauses lays bare what gaps exist in the Maltese relativisation system and how these gaps get circumvented via other means in the grammar. It further allows us to better evaluate certain behaviours whilst pinpointing what additional work still needs to be done on the subject.

Dan ix-xogħol jiġbor fil-qosor il-letteratura li fiha tissemma s-sentenza subordinata aġġettivali (SSA) fil-Malti u juri li s-SSA fil-lingwa mhijiex struttura omoġenja. Jiġu diskussi tliet tipi ta'

SSA: restrittivi, mhux restrittivi u ħielsa. Dawn iġibu magħhom restrizzjonijiet individwali kemm fuq l-anteċedent (meta jkun preżenti) kif ukoll fuq l-istrateġiji differenti li jużaw. Toħroġ ċara d-distinzjoni bejn is-sintassi tal-Malti Standard u tad-djalett fl-użu tal-istrateġija pronominali, għall-inqas fi strutturi tas-SSA mhux ħielsa. Id-diskussjoni turi wkoll kif id-disponibbiltà tas-SSA ħielsa li għandhom konġunzjoni subordinata fil-Malti hija tipoloġikamant rari. Din il-ħarsa ġenerali lejn dak li nafu dwar is-SSA fil-Malti tesponi l-lakuni li hemm fis-sistema u turi kif dawn il-lakuni jiġu evitati bis-saħħa ta' mezzi oħra grammatikali. Barra minn hekk, inkunu f'qagħda aħjar li nevalwaw xi mġiba lingwistika u nagħrfu x'jista' jsir aktar fuq is-suġġett.

1. Introduction

The (morpho)syntax and semantics of different relative clause types in Maltese have recently received quite some attention. The presentation of this work here highlights the insights provided in Camilleri (2012), Camilleri (2014a), Camilleri & Sadler (2011), Camilleri & Sadler (2012a), Camilleri & Sadler (2016), Sadler & Camilleri (2018), rectifying, and sharpening the claims made therein. Here I choose to concentrate on three broad types of relative clauses (RCs) in Maltese, namely, restrictive relative clauses (RRCs), non-restrictive relative clauses (NRRCs), and free relative clauses (FRCs). I discuss the structure that constitutes the distinct type of clauses, the strategies employed in the expression of the different functions associated with the different RCs, and the constraints that govern the morphosyntactic interface to yield different semantic readings, which also includes reference to the strategies employed internal to the relative clause itself. The paper proceeds as follows. First I establish the major differences that characterise the different RCs under investigation (§2), and in §3, in what is the bulk of the study, I concentrate on the landscape of

strategies employed to introduce Maltese RCs, and the constraints that govern them. §4 provides a parenthesis that specifically focuses on FRCs, as particularly befits the discovery of a type of FRC in the grammar, which has been shown to be quite rare, crosslinguistically, while §5 concludes with the insights of this study.

2. A divide in form and function

The relative clause (RC) more broadly functions as a means with which to add information and elaborate upon a referent, the antecedent, which can be known, or otherwise, in which case, reference becomes identified via the presence of the RC. In (1), *the food* (i.e. the antecedent) being referred to is specifically the one that has been cooked for her, as opposed to any other *food* supply that may be available, or known from within the discourse context. Crucially, the antecedent bears a function, within the RC. In (1), *the food* functions as the direct object.

(1) *the food that they've cooked for her*

The structure of a RC is construed as involving a nominal antecedent, and an adjunct clause. Specifying here that the clause involved functions as an adjunct clause precludes the possibility of an alternative analysis that considers the clause as some complement to the nominal antecedent, as is the case with factual clauses of the type *the fact that*, in which the *that* clause is a complement of *the fact*. Specifying that the antecedent takes an in-clause function in turn excludes structures such as *why he came* in *the reason why he came in* from being interpreted as RCs. The above characterisation constitutes the prototypical structure true of both RRC (such as (2a)), and NRRC (2b) type constructions.

(2a) *I will eat the food*$_{[antecedent]}$ *{that/which they'll give me}*$_{adjunct\ clause}$. RRC

(2b) *I will eat the food*$_{[antecedent]}$, *{which they gave me}*$_{adjunct\ clause}$. NRRC

The structural characterisation that is true of RRCs and NRRCs does not hold for FRCs. In the literature, a number of labels have been used to refer to such types of RCs. Huddleston & Pullum (2002) use the term *fused* relative clauses, indicative of the fusion of the nominal antecedent and the *wh*-pronoun used to characterised English FRCs. Another term is *headless*, which is the one employed in the descriptive grammar of Maltese in Borg & Azzopardi- Alexander (1997). This terminology is usually laden with analytical concerns which we do not need to delve into, here (see e.g. Grosu & Landman (1998), Izvorski (2000), Citko (2002) for related discussions). The use of the term *headless* aligns with an analysis that views FRCs as void of an overt head, given that, as illustrated in (3), in contrast to the structures in (2), there is no distinguishable antecedent.

(3) *I will eat what they'll give me.*

In (3), as opposed to (2), there is no identifiable NP that can be said to function as the antecedent, and which is separate, or distinct from the *wh*-pronoun introducing the clause which modifies that antecedent. In contrast, what we have here is 'just' a clause, with the *wh*- pronoun *what* which 'doubles' its function both as the (nominal and non-clausal) argument of the (matrix) predicate *eat*, and a clause which additionally functions as the modifier of the same incorporated fused argument.

Having established broadly the major structural (and formal) difference between RRCs/NRRCs, on the one hand, and FRCs, on the other, we now focus on the semantic differences which obtain with respect to the function expressed by the adjunct clause part of the RC construction. The function of the RRC (as is also in essence that of a FRC but in perhaps a more opaque manner) is to

act as an intersective modifier that is meant to further specify (and identify) the antecedent.[1] As its name suggests, the function of this type of clause is to restrict the antecedent's reference. In contrast, a NRRC's function is to add more information about, or to elaborate upon whatever property is to be associated with the antecedent. This entails, in turn, that the antecedent of the NRRC involves an already specified entity that has been/is anchored contextually, or otherwise, e.g. via shared knowledge between the interlocutors, in the discourse interaction. Huddleston & Pullum (2002), for instance, refer to NRRCs with the label *supplementary relatives* whose function is to add and contribute further to some already known knowledge. Therefore, the NRRC, as opposed to the RRC is not meant to distinguish its antecedent from other members within a set. The contrastive reading that obtains between the choice of one RC as opposed to the other can be observed through the pair in (4), whereas per convention, the NRRC is distinguished from a RRC by means of a comma (,) that comes in between the antecedent and the clause. (3a) clearly identifies the book that was bought as being the cheapest book member out of a set of non-paperback books, while the function of the NRRC in (3b) is to add more information about the nature of the cheapest book bought, which, as it happens, is not a paperback. Further evidence that the semantics of the RRC is mainly to restrict reference can be illustrated by the substitution of the RRC by an attributive adjective. (3a) can thus read as (5).

(4a) *I bought the cheapest book {which was not a paperback}.* RRC

(4b) *I bought the cheapest book, {which was not a paperback}.* NRRC
Arnold (2007, p. 272)

(5) *I bought the cheapest non-paperback book.*

1 An intersective modifier is a type of adjective that does not change the category of the noun in question, and its content remains true independent of what it combines with.

Parallels to (4) obtain in the representative data set for Maltese in (6).

(6a)	*It-tifla*	*{li n-af*	*jien}*	*kellm-it-ni*	*lbieraħ.*
	DEF-girl	LI 1-know.IPFV.SG	I	talk.PFV-3SGF-1SG.ACC	yesterday
	'The girl that I know talked to me yesterday.'[2]				RRC

(6b)	*It-tifla,*	*{li għad-ha*		*kemm*	*ġiet*
	def-girl	LI just-3SGF.GEN		how much	come.PFV-3SGF
	t-kellim-ni},		*qal-t-l-i*		*li ...*
	3F-talk.IPFV.SG-1SG.ACC		say.PFV.3-SGF-DAT-1SG		COMP
	'The girl, who has just come to talk to me, told me that ...'				NRRC

In (6a), the antecedent *it-tifla* 'the girl' is identified from the larger set of girls in which it participates as a member. In contrast, the function of the adjunct clause as part of the larger NRRC structure in (6b) is merely to add more information about some already-anchored antecedent.

Concomitant with the distinct semantic characteristics that differentiate RRCs from NRRCs are syntactic constraints that have to do with the order of the RCs vis-à-vis one another, when they co-occur. It is possible to have the same RC type co-occurring, as illustrated through (7a), which involves the stacking of two NRRCs. The same follows for RRCs. In contrast, a general linear ordering constraint holds when two RCs that are not of the same type co-occur; a RRC (or FRC) must precede a NRRC, as illustrated in (7b). The obligatory requirement of the RRC to linearly precede the NRRC follows from the distinct semantic function of the two types of adjunct clauses, where the RRC's function in structures involving stacked RCs is to initially restrict (fully) and anchor the reference of the antecedent. The NRRC that follows, then takes to the task to provide additional information about the already established reference.

2 For now, I will just gloss *li* as LI so as not to engage in an analysis of this item, as yet.

(7a) *Mario, li n-af jien, li dejjem*
Mario, LI 1-know.IPFV.SG I, LI always
i-dur waħd-u l-bandl-i, ...
3M-go around.IPFV.SG alone-3SGM.GEN DEF-swing-PL
'Mario, whom I know, who is always going around alone in the playing field ...'
NRRC + NRRC

(7b) *It-tifel li n-af jien, li dejjem*
DEF-boy LI 1-know.IPFV.SG I, LI always
i-dur waħd-u l-bandl-i ...
3M-go around.IPFV.SG alone-3SGM.GEN DEF-swing-PL
'The boy who I know, who is always going around on his own in the playing field ...'
RRC > NRCC

Beyond considerations that have to do with ordering and co-occurrence constraints, RRCs and NRRCs are additionally differentiated on the basis of the constraints they are subject to, with respect to the antecedents they are able to modify. Below in Table (1) is a list of distinct antecedents along with a reference to their ability (or otherwise) to function as antecedents of a RRC or NRRC, or both. The data in (8)-(10) are then meant to illustrate several of these types of antecedents and the RC they are able to occur with.

Antecedent type	**RRC**	**NRRC**
NP: *tifel*/*it-tifel* '(the) boy'	✓	✓
temporal NP: *il-ġimgħa d-dieħla* 'the next week'	*	✓
Proper Name: *Marija* 'Mary'	*	✓
il-Proper Name: *il-Marija* 'the Mary'	✓	*
free pronoun: *jien* 'I', *lilek* 'you.NON-NOM'	✓	✓
clausal	*	✓
negative universal quantifier: *ebda* 'no(ne)'	✓	*
positive quantifier: *uħud* 'some', *kollha* 'all', *kull* 'every'	✓	✓
negative universal NP: ħadd 'no one', *xejn* 'nothing', *mkien* 'nowhere'	✓	*
positive universal NP: *kulħadd* 'everyone', *kollox* 'everything', *kullimkien* 'everywhere'	*	*
split antecedent	*	✓

Table 1: Constraints on the antecedent types available when comparing RRCs vs. NRRCs

The contrast in (8), for instance, brings out the differences in the intricacies associated with Proper names as antecedents in the context of RRCs vs. NRRCs. The use of the definite article in (8b) is indicative of the fact that the antecedent *Mario* is being identified from a set of referents called *Mario*. The RRC puts the specific entity *Mario* in contrast with other referents that are also called *Mario*.[3]

(8a) *Mario,* *li* *daħal* *issa …*
Mario LI enter.PFV.3SGM now
'Mario, who's entered now …' NRRC

(8b) *Il-Mario* *li* *daħal* *issa …*
DEF-Mario LI enter.PFV.3SGM now
'The Mario who's entered now ...'
RRC: Camilleri & Sadler (2016, 118)

(9) shows RCs with clausal antecedents which, as represented in Table (1), can only appear in the context of a NRRC.

(9a) *[Marija* *poġġie-t* *kollox* *f'kamrit-ha],* *li*
Mary place.PFV-3SG everything in.room.SGF-3SGF.GEN LI
fil-verità *kien* *l-aħjar* *li* *setgħ-et*
in.DEF-truth be.PFV.3SGM DEF-good.ELAT LI can.PFV-3SGF
t-a-għmel.
3F-FRM.VWL-do.IPFV.SG
'Mary placed everything in her room, which in reality was the best thing she could have done.'

(9b) *Imbagħad* *[Kim* *beda* *j-suq*
then Kim start.PFV.3SGM 3M-drive.IPFV.SG
bl-addoċċ], *li* *fil-fatt* *huwa*
with.DEF-random LI in.DEF-fact COP.3SGM
perikoluż *ħafna.*
dangerous.SGM a lot
'Then Kim started to drive haphazardly, which is indeed very dangerous.'
NRRC: Camilleri & Sadler (2016, p. 121)

3 It should perhaps be mentioned here that at times, especially in colloquial speech, the Proper Name can easily function as an antecedent of a RRC without the need to mark that Proper Name as [+DEF] via the presence of the article. The antecedent of this type of RRC implies that the speaker-hearer happen to have multiple common referents that share the same name.

The contrastive data in (10) is meant to display the differences that obtain in the context of negative vs. positive universal indefinites as antecedents, in particular. The major difference, as also represented in Table (1), is the fact that while negative universal indefinites *can* function as antecedents, even if restricted to RRCs, as illustrated through (10a-10b), positive counterparts cannot function as antecedents, as the ungrammaticality of (10c-10d) illustrates, creating an interesting POLARITY-based split in the grammar.

(10a) *Ma kien hemm ħadd li ma*
NEG be.PFV.3SGM EXIST no one.SGM LI NEG
kon-t-x n-af-u qabel.
be.PFV-1SG-NEG 1-know.IPFV.SG-3SGM.ACC before
'There was no one that I didn't know before.' RRC

(10b) ***Xejn**, li x<t>aq-et t-i-sma',*
nothing.SGM LI wish.REFL.PFV-3SGF 3F-FRM.VWL-hear.IPFV.SG
ma nt-qal.
NEG PASS-say.PFV.3SGM
'*Nothing, which she wanted to say, was said.'[4] *NRRC

(10c) ***kulħadd**(,) li mar ...*
everyone.SGM LI go.PFV.3SGM
Intended: 'everyone that went…' *RRC/NRRC

(10d) ***kollox**(,) li għid-t-l-ek ...*
everything.SGM LI say.PFV-1SG-DAT-2SG
Intended: 'all that I told you …' *RRC/NRRC

It is quite interesting that the observed gap in the context of positive universals, as displayed in their inability to function as

4 The gloss FRM.VWL in relation to the *i* in the imperfective form *tisma'* refers to the formative vowel (Puech, 1979) that comes in between the prefix and the stem in the imperfective sub-paradigm, and similarly, the vowel that precedes the stem in the imperative sub-paradigm. It is essentially an arbitrary morphological form that functions as a phonological extension of the morphological stem in the imperfective and imperative sub-paradigms, and which is conditioned, or governed by phonological constraints. Refer to Camilleri (2014b), for more details.

antecedents in the context of RRCs, which is where the contrast with respect to negative indefinites holds, is a reflex of yet another POLARITY-based split in the grammar. Maltese displays positive universal *wh*-pronouns, but lacks negative counterparts. For this reason, the gap that results, as evinced through the ungrammaticality of (10c-10d), is made up for by means of a FRCs, which, as mentioned above, are semantically closer to RRCs than NRRCs. The FRC data that in Maltese is used to substitute the ungrammaticality of positive universal indefinites-headed RRCs is provided below in (11). Similarly, the inability of the positive universal indefinite *kullimkien* 'everywhere' to function as the antecedent of a RRC is made up for by the use of the *wh*-pronoun *kull fejn* 'everywhere', in a FRC context, as in (12).

(11a) ***kull min*** *mar* ...
whoever go.PFV.3SGM
'whoever went …'

(11b) ***kulma*** *għid-t-l-ek* ...
whatever say.PFV-1SG-DAT-2SG
'Whatever I told you …' FRC

(12) ***kull fejn*** *t-mur*
wherever 2-go.IPFV.SG
'wherever you go' FRC

Beyond the nature of the constraints on the antecedent, and the actual function of the different RCs, yet another difference which distinguishes RRCs from NRRCs is the **head parameter**, i.e. the parameter that has to do with where the antecedent linearly occurs, vis-à-vis the adjunct clause. While RRCs in Maltese are always externally-headed, as illustrated through (5a) and (6b) above, for instance, i.e. where the antecedent sits outside of the RC proper, specifically at the left-edge, given the language's head-initial parameter, NRRCs in Maltese, on the other hand, can be of two types. They can be either externally-headed, as

observed through the array of different NRRC examples above, or internally-headed, even if rather constrained, when so. Internally-headed NRRCs in Maltese require the concurrent presence of an external head, with anaphoric-binding occurring between the two. A parallel constraint, which only applies for, and is restricted to NRRCs, as opposed to RRCs, is also found in English and Italian. An instance from the latter is in (13), where the internal NP *romanzo* 'novel' is co-indexed (marked via the subscript *i*) with the external antecedent of the construction.

(13) *Ha raggiunto la fama con [Il giardino*
has reached DEF.SGF fame.SGF with DEF.SGM garden.SGM
*dei finzi-contini]$_i$ {**il quale romanzo**$_i$*
of.PL F-C DEF.SGM which novel.SGM
ha poi anche ... }
has also even
'He became famous with *Il giardino dei Finzi-Contini*, which novel was then also ...'
Italian: Cinque (2008, p. 106)

In English, examples of such internally-headed structures include (14). (14a) involves the internal head *society* co-indexed with *the LAGB*, while (14b) is somewhat more complex, where the internal head is in fact co-indexed specifically with the quantifier/numeral that gov- erns, modifies or specifies (depending on one's theoretical analysis) the RC's antecedent.

(14a) *[The LAGB]$_i$ {**which society**$_i$ was founded in ...}* (Citko, 2008, p. 635)

(14b) *There were only [[thirteen]$_i$ senators] present, {**which number**$_i$ was too few for a quorum}.* (Arnold, 2007, p. 289)

The Maltese internally-headed NRRC data is just as interesting. Beyond a clear demonstration of the fact that this structure is available, as shown through the data in (15), Maltese introduces internally-headed NRRCs with a very particular item: the *wh*-pronoun *liema* 'which'. In (15a), for instance, we observe the

internal head's co-indexation with a coordinated set of antecedents that head the relative clause. In (15b) we get to observe how, and in which way, it becomes possible to have **split antecedents** in the context of NRRCs (but not with RRCs, as represented in Table (1)). The internal-head *frott* 'fruit' gets co-indexed with two antecedents that are present in two distinct clauses. The data in (15c) is there to additionally exemplify that it is possible to also have an internal-head embedded within a pied-piped relative clause, i.e. one in which the clause functions as a complement of a preposition (*fi* 'in' in this case), heading a PP (*f'liema post* 'in which place'), which appears in a fronted position to the left-edge of the RC.

(15a) *{Pawlu* *u* *Salvu}$_i$,* *liema* *rġiel$_i$* *qal-u* *li* ...
Paul CONJ Salvu, which men say.PFV.3-PL COMP
'Paul and Salvu, which men said that …'

(15b) *Marija* *t-ħobb* ***it-tuffieħ$_i$*** *filwaqt* *li*
Mary 3F-love.IPFV.SG DEF-apple.MASS while COMP
Rita *t-ħobb* ***il-banana$_j$*** ***{liema frott$_{i+j}$***
Rita 3F-love.IPMV.SG DEF-banana, which fruit.MASS
dejjem *j-eħd-u-h* *magħhom* *għal-lunch}.*
always 3-take.IPFV-PL-3SGM.ACC with-3PL.GEN for.DEF-lunch
'Mary loves apples, while Rita loves banana, which fruit they always take with them for lunch.'
Camilleri & Sadler (2016, p. 121)

(15c) *Il-Palazz$_i$* ***f'liema post$_i$*** *t-laqqgħ-u*
DEF-palace in.which place PASS-CAUSE.gather.PFV.3-PL
l-mistedn-in ...
DEF-guest-PL
'The Palace, in which place the guests where gathered …'
Camilleri & Sadler (2012a, p. 20)

A clause introduced by *liema* is not the only strategy used in Maltese to express internally-headed NRRCs. *Liema* functions as some sort of specifier to the internal head. However, when the internal head is not specified via what is termed as a *wh*-pronoun in English, such that it is required to be specified via

other means, such as syntactic/analytic adjunction, it is possible to resort to the usual *li*, which has characterised the RRC/NRRC structures prior to the discussion associated with considerations of an internal head. An illustration of the employment of *li* in a context involving an internal-head is (16). In the presence of *li*, the internal head is specified via the adjunction of the PP *bħal din* 'like this.SGF'. This specification then co-indexes the internal NP head *ħaġa* with the antecedent external to the clause.

It is needless to say that, the *liema* strategy would have worked just the same here, and it is only for reasons of space that I am not providing an example of the alternative. The employment of *liema* can thus be understood as being in a complementary distribution with the use of *li* + PP adjunction. *Liema* would thus be possible only with a non-PP modified *ħaġa* as the internal head. What is further special with the internally-headed NRRC in (16) is the fact that it shows how co-indexation does not necessarily imply agreement resolution, as is the case of (15a-b), or agreement matching, as in the case of (15c). Rather, while the RC's antecedent's head is *qtil* 'killing.SGM', the internal head is *ħaġa* 'thing.SGF' in (16).

(16) *[Il-qtil* *tat-tifel]*$_i$ ***li [ħaġa*** ***bħal*** ***din]***$_i$
DEF-killing.SGM of.DEF-boy LI thing.SGF like DEM.SGF
ma *stennej-nie-ha*$_i$ *qatt,* *ħasad*
NEG expect.PFV-1PL-3SGF.ACC never shock.PFV.3SGM
lil *kulħadd.*
ACC everyone
'The boy's killing, which was something no one expected, shocked everyone.'
Agreement mismatch: Camilleri & Sadler (2012a, p. 25)

With this discussion of the core differences in the form and function of RRCs and NRRCs in particular, we now move on to consider the relativisation strategies available for Maltese RCs, which has been something I briefly touched upon in the last part of the discussion on internally-headed RCs when contrasting the

constraints that govern the complementary distribution of the *wh*-specifier *liema* and the use of *li* in the context of internally-headed RCs. Further discussion follows below.

3. Relativisation strategies

As established in Camilleri (2010), the strategies employed to introduce Maltese RCs are: (i) the use of *li*, (ii) the use of a *wh*-pronoun (as is the case of *liema* in the context of internally-headed NRRCs discussed earlier, for instance), and (iii) a ø (zero) strategy. Notwithstanding the availability of the *wh*-pronoun strategy in Maltese, it is constrained in quite a complex way in Standard Maltese, as opposed to the laxer distribution it displays in non-Standard varieties (see Camilleri (2012) for more detail). In line with Fabri (1987), Borg (1991), Borg (1994), Borg & Azzopardi-Alexander (1997), I analyse *li* as a complementiser (see Camilleri (2014a), Camilleri & Sadler (2016), Sadler & Camilleri (2018) for more detail), in contrast to Sutcliffe (1936) and Aquilina (1973), who analyse it as a *wh*-pronoun.[5] While the former two strategies can introduce both RRCs and NRRCs (as well as FRCs, as we will see in §4), the ø zero strategy is highly constrained, and is additionally almost exclusive to RRCs.[6] An illustration of the *wh*-

5 From now onwards I will thus be glossing *li* as COMP, indicative of the complementiser (C) category and consequently the C position I consider this item to take at the constituent-structure level.

6 That distinct strategies are employed in the context of different RC types, or that restrictions are imposed upon the array of strategies involved, or the extent of their employment, is not something that occurs only in Maltese. In English, for instance, NRRCs can only be introduced through the *wh*-pronoun strategy, in contrast to the *that* or zero strategies which are additionally able to introduce RRCs. In Italian too, for instance, the *wh*-pronoun strategy involving *il quale* is used instead of *cui*/*che* in the context of NRRC (as exemplified in (13) above). Moreover, English NRRCs and FRCs, for instance, which are obligatorily introduced by a *wh*-pronoun do not involve the same set of pronouns. *Ever*-type *wh*-pronouns, such as *whoever*, *whenever*, and others, are only available as a subset of the *wh*-pronouns that can introduce FRCs.

pronoun strategy, and the ø (zero) strategy is provided through the data in (17).

(17a) *it-tifel* *ma'* ***min*** *kon-t*
DEF-boy with who be.PFV-1SG
'the boy with whom I was'[7] *wh*-pronoun strategy

(17b) *Tifel* ø *j-o-qtol* *il-qtates* *mhux* *se*
boy 3M-FRM.VWL-kill.IPFV.SG DEF-cat.PL NEG PROSP
j-i-bża' *minn* *ġurdien.*
3M-FRM.VWL-fear.IPFV.SG from mouse
'A boy who kills cats is not going to fear a mouse.'
ø strategy: Borg & Azzopardi-Alexander (1997, p. 35)

Apart from *li* as a complementiser, Maltese also employs *milli* (see *milli*), which in Camilleri (2010) and subsequent works is referred to as a partitive complementiser, at least in its function to introduce RCs. Its partitive function is clearly carried forward from the fusion of the P *minn* 'from' along with the complementiser *(il)li*, which, in turn, provides the antecedent with an element out of a set reading. This then impinges on the nature of the antecedent, which must be indefinite. *Milli* as a complementiser in the grammar does not solely occur in the context of partitive RCs of the type in (18). Rather, *milli* also functions as a complementiser that introduces an adjunct clause at the sentential level, meaning 'from, instead of', rather than solely an adjunct clause at the NP level, as in the case of RCs. This function is exemplified through the Maltese proverb in (19).

(18) *Għoġb-ok* *xi* *ktieb* ***milli***
like.PFV.3SGM-2SG.ACC some book from.COMP
ġib-t-l-ek?
bring.PFV-1SG-DAT-2SG
'Did you like any book from the ones that I brought you?'

7 Note that this structure may not be acceptable for all, and may be indicative of dialectal variation.

(19) *Aħjar* *uff* ***milli*** *aħħ.*
good.ELAT uff from.COMP ahh
Lit: It is better an uff, instead of an ahh.
It is better to complain for a while, instead of suffering, or feeling pain.
Maltese Proverb

3.1 The *wh*-pronoun strategy

While perhaps the *li* strategy is the most commonly used in Maltese to introduce RRCs and NRRCs (and FRCs (§4)), together with *milli*, which is less common, *wh*-pronoun introduced RCs have not been studied much. Borg & Azzopardi-Alexander (1997) only discuss them with respect to what we are here referring to as FRCs.[8] As illustrated through (17a), however, non-FRCs can indeed be introduced by a *wh*-pronoun strategy in Maltese. We have in fact in §2 seen that internally-headed NRRCs can also be similarly-introduced in Maltese.

Focusing on Standard Maltese, the *wh*-strategy is widely used in pied-piped contexts. (17a) above is one such instance. It involves the use of the [+HUMAN] *wh*-pronoun *min* 'who', while (20a) below involves the use of the [-HUMAN] *wh*-pronoun counterpart *xiex* 'what'.[9] Such pied-piping contexts can easily be substituted by the *li* strategy, as in (20b), along with changes in the morphosyntax internal to the RC, to be discussed further below.

(20a) *It-trav-i* *{ma'* ***xiex*** *j-i-d-dendl-u*
DEF-beam-PL with what 3-EPENT.VWL-REFL-hand.IPFV-PL
l-qniepen}, *is-sadd-u.*
DEF-bell.PL REFL-rust.PFV.3-PL
'The beams to which the bells are hung, have got rusted.' MLRS

8 I am here deliberately excluding reference to Aquilina (1973), for instance, who treats *li* as a *wh*-pronoun. Moreover, similar to Borg & Azzopardi-Alexander (1997), Sutcliffe (1936) only discusses *wh*-pronouns in the context of FRC structures.

9 Yet again, one should mention that this structure may not necessarily be considered grammatical in Standard Maltese, even if it constitutes part of the *MLRS* Corpus.

(20b) *It-trav-i*$_i$ *{**li*** *j-i-d-dendl-u*
DEF-beam-PL COMP 3-EPENT.VWL-REFL-hang.IPFV-PL
magħ-hom *il-qniepen},* *is-sadd-u.*
with-3PL.GEN DEF-bell.PL REFL-rust.PFV.3-PL
'the beams that the bells are hung on to, have got rusted.'
Camilleri (2014a, p. 185)

Further evidence indicative of the widespread use of a *wh*-pronoun strategy in the context of RRCs/NRRCs, particularly within pied-piped structures, comes from the grammaticalisation of new *wh*-pronouns in the grammar that have come about via the univerbation, i.e. the fusion of two distinct and separate word-forms, which in this case are a P and a *wh*-pronoun; parallel to the process that renders the complementiser *milli* just referred above. This process is also suggestive of the linear adjacency that governed the P and *wh*-pronoun items prior to their fusion, which would have in turn also been precisely what facilitated, and led to the very fusion. Such univerbated *wh*-pronominal instances include *fiex* (< *fi* 'in' + *xiex* 'what') and *mnejn* (< *minn* 'from' + *fejn* 'where'), as represented through (21a) and (21b), respectively. (Refer also to Table (2) below).

(21a) *Xtraj-t* *kaxxa* ***{fiex*** *in-qegħid-hom}.*
buy.PFV-1SG box in what 1-place.IPFV.SG-3PL.ACC
'I bought a box to put them in.'

(21b) *T-af-ha* *t-triq* ***{mnejn*** *t-i-sta'*
2-know.IPFV.3SG.ACC DEF-road.SGF whence 2-EPENT.VWL-can.IPFV.SG
t-għaddi}.
2-pass.IPFV.SG
'You know the way from where you can pass.'

In the Standard variety, beyond the use of a *wh*-pronominal strategy in pied-piping contexts, antecedents that take a locative thematic-role can be similarly introduced. In such instances, it is the *wh*-pronoun *fejn* 'where' that is employed, as in (22) below. Once again, this is optional, as the *li* strategy along with concurrent morphosyntactic changes can also be employed.

(22) *Ir-restorant* ***fejn*** *mor-na* *d-darba* *l-oħr-a* ...
DEF-restaurant where go.PFV-1PL DEF-once.SGF DEF-other-SGF
'The restaurant where we went last time ...'

From the above characterisation of the constraints that govern the *wh*-pronominal strategy in the Standard variety, we appear to have a strategy that is 'restricted' to contexts involving antecedents that take an adjunct (ADJ) function, and an oblique (OBL) and oblique object (OBL OBJ) (i.e. object argument of a preposition) in-clause grammatical function, i.e. the NP which the antecedent displays a dependency on, internal to the RC. I use the term 'restricted' here in the context of Keenan & Comrie's (1977) Accessibility Hierarchy, presented in (23) below.

(23) SUBJ > DO > IO > OBL > GEN (possessor) > OCOMP (object of comparison)
Accessibility Hierarchy: Keenan & Comrie (1977, p. 66)

The hierarchy should be interpreted such that the grammatical function furthest on the left-edge is understood to be more accessible for relativisation than the one that follows it on the right-edge, crosslinguistically. Hence, if a grammatical function lower on the hierarchy is available for relativisation in a particular linguistic system, then the expectation is such that any grammatical function higher on the hierarchy, i.e. to the left, would be also available for relativisation. While I will not engage in a discussion as to how much more fine-grained the grammatical functions on the Accessibility Hierarchy ought to be for Maltese (see Camilleri (2010), Camilleri (2014a), Camilleri & Sadler (2016) for more detail), what is key to our observation at this point in the discussion is the fact that the *wh*-pronoun strategy in Standard Maltese appears to be *unusually* confined to positions lower on the hierarchy. From the distribution as laid out above, these positions specifically include obliques and object of prepositions, as well as adjuncts, which would be positioned lower still, on the hierarchy in (23), given that adjuncts are not subcategorised arguments as the

rest of the grammatical functions on the Accessibility Hierarchy are.

This landscape is in contrast with the use of the *wh*-pronoun strategy in dialectal Maltese, where it can additionally be used with [+HUMAN] antecedents that display a dependency with in-clause functions other than the ones just listed above for the Standard variety. These include direct, and indirect object functions (i.e. OBJ and OBJθ, respectively). The latter is the case in (24). A constraint appears to hold, however; the antecedent of such RCs is constrained to be [+DEF], (apart from being [+HUMAN]).

(24) *Ilbieraħ,* **(ir)-raġel* ***'il*** ***min*** ċempil-t,
yesterday DEF-man DAT who phone.PFV-1SG
qdie-ni.
serve.PFV.3SGM-1SG.ACC
'Yesterday, the man whom I phoned attended to me.'

While the dialectal scenario provides us with a wider distribution of the *wh*-pronominal RC strategy, in comparison with the Standard variety, a glaring gap remains in the system, and that is the absence of [-HUMAN] RC antecedents, whether definite, or otherwise. Constructions such as (25) are ungrammatical, even if the *xi*/*x'* 'what' [-HUMAN] *wh*-pronoun presents itself as an available counterpart to [+HUMAN] *min* 'who' in the grammar.

(25a) **l-aħbar* ***x'għaġġb-et*** *lil* *kulħadd*
DEF-news.SGF what.surprise.PFV-3SGF ACC everyone
Intended: 'the news that surprised everyone' Camilleri & Sadler (2016, p. 120)

(25b) **Xtraj-t* *ktieb* ***xi*** *n-sellef.*
buy.PFV-1SG book what 1-lend.IPFV.SG
Intended: 'I bought a book to be able to lend.'

Notwithstanding the ungrammaticality of the above examples, it turns out, however, that the gap associated with the absence

of [-HUMAN] antecedents in the system is *only* apparent. At first sight, it *does* translate as a gap, just as (wrongly) claimed in the earlier works in Camilleri (2010), Camilleri (2012), and Camilleri & Sadler (2011); however, this is only because of the highly constrained nature of the structure that can allow for the use of *xi*/*x'* in both Standard and dialectal RCs. On unravelling this possibility in the system, the [+HUMAN] counterpart, which then makes use of the *wh*-pronouns *min* 'who'/*'l min* 'whom', also becomes available to the Standard variety, so long as it is governed by the same set of constraints.

Camilleri (2014) identifies the following set of constraints said to govern the availability of *xi*/*x'* in Standard Maltese, with the final constraint having been identified later in Camilleri & Sadler (2016), and then discussed and developed further in Sadler & Camilleri (2018). It was also in the latter works that it also became clear that this same set of constraints also governs the use of *min* 'who' in the Standard variety, beyond its uses in association with adjunct, oblique, and object of preposition in-clause functions.

1. [-DEF] (indefinite) antecedent;
2. matrix clause function of the antecedent can only be a term, particularly a SUBJ, OBJ, or OBJ *theme* (i.e. non-DAT);
3. in-clause function can only be a term of the type: SUBJ, OBJ, or OBJθ (i.e. DAT/non- DAT);
4. imperfective RC predicate (excluding any ASPECTUAL augmentation via auxiliaries);
5. the matrix predicate must entail an existential component in its semantics, expressing notions of coming into being, view, or availability via possession or transfer, and the like.[10]

10 It is this lexical dimension that pertains to the predicates that take such indefinite-headed RCs as their argument, that the literature refers to these

Examples instantiating this set of constraints is provided through the data in (26) below. In (26a), the indefinite NEG universal antecedent *xejn* 'nothing' is the OBJ of the possessive predicate in the matrix that then functions as the SUBJ of the RC's (imperfective) predicate *dejjaq* 'bother'. In (26b), the indefinite *rota* 'bicycle' functions as the OBJ in both the matrix clause, headed by the stative *fadal* 'remain', and the RC. In contrast, the (quantified) indefinite antecedent *ħobż* 'bread' is the OBJ theme of the distransitive matrix (transfer-of-possession) predicate *ta* 'give', which is then in a dependency with the OBJ of the verb *xewa* 'toast' within the RC. (26d) provides us with an illustration of the antecedent functioning both as the SUBJ of the matrix predicate, as well as the SUBJ of the RC's (imperfective) predicate.

(26a) *M'għand-i* *xejn* ***{xi***
NEG.have-1SG.GEN nothing.SGM what
j-dejjaq-ni}.
3M-bother.IPFV.SGM-1SG.ACC
'I have nothing that's bother me.' Sutcliffe (1936, p. 182)

(26b) *Fadal* *rota* ***{xi*** *n-ġib}*.
remain.PFV.3SGM bicycle.SGF what 1-get.IPFV.SG
'There remains a bicycle to bring along.'

(26c) *Ta-ni* *biċċt-ejn* *ħobż*
give.PFV.3SGM-1SG.ACC piece.F-DU bread
{x'n-i-xwi-l-hom}.
what.1-FRM.VWL-toast.IPFV.SG-DAT-3PL
'He gave me two pieces of bread to toast for them.'

(26d) *J-eżist-u* *alternattiv-i* *oħr-ajn* ***{x'j-i-stgħ-u***
3-exist.IPFV-PL alternative-PL other-PL what.3-EPENT.VWL-can.IPFV-PL
j-i-nt-użа-w}.
3-EPENT.VWL-PASS-use.IPFV-PL
'There exist other alternatives that can be used.'

types of RCs as *headed modal existential constructions*. In §4 we will consider the non-headed counterparts.

The use of *min* in the [+HUMAN] counterpart is illustrated below. In (27), the indefinite antecedent *xi ħadd* 'someone' functions as the internal argument of the existential predicate *hemm*, and displays a dependency with the OBJ of the predicate *kellem* 'talk'.

(27)	*Hemm*	*xi*	*ħadd*	***{'il***	***min***
	EXIST	some	no one	ACC	who
	n-i-stgħ-u			*n-kellm-u}?*	
	1-EPENT.VWL-can.IPFV-PL			1-talk.IPFV-PL	
	'Is there anyone whom we can talk to?'				

(27) in the Standard variety thus stands in contrast to the lesser constrained distribution of min 'who' in the dialect, where, as illustrated through (24) above, can also be employed in the context of [+DEF] antecedents. It is however interesting to observe that a gap remains in the unavailability to relativise [-HUMAN] [+DEF] antecedents in both the Standard and non- Standard varieties.

Table (2) summarises the facts, and brings in one place the rich array of *wh*-pronouns that can introduce RRCs and NRRCs in Maltese.

3.2 The zero strategy

While it would be possibly fair to say that the zero (ø) strategy is the least widely distributed, it is also the most constrained. If we maintain our focus on finite RCs, rather than considering RCs involving participial forms, then RCs introduced by a zero strategy are constrained to involve:

1. [-DEF] antecedent;
2. imperfective RC predicate, if the construction is verbal;
3. in-clause function can only be an immediate- or long-distance SUBJ or POSS

Antecedents - in-clause function	***wh*-prn**
[+HUMAN] [+DEF] - OBJ/OBJθ	*min* 'who'
[+HUMAN] [+DEF] - OBJ/OBJθ	*'l min* 'whom'
[+HUMAN] [-DEF] & matrix OBJ/OBJ *theme* - SUBJ/OBJ/OBJT	*'l min* 'whom'
[-HUMAN] [+DEF]	n.a
[-HUMAN] [-DEF] & matrix OBJ/OBJ *theme* - SUBJ/OBJ/OBJθ	*xi; x'* 'what'
[+HUMAN] - OBL OBJ	P + *min*
[-HUMAN] - OBL/ADJ	*fuqiex* 'on what' < *fuq* 'on' + *xiex* 'what' *fiex* 'in what' < *fi* 'in' + *xiex* 'what' *biex* 'with what' < *bi* 'with' + *xiex* 'what' *mniex* 'from what' < *minn* 'from' + *xiex* 'what' *għaliex* 'for what' < *għal* 'for' + *xiex* 'what'
[-HUMAN] - OBL OBJ/ADJ OBJ	P + *xiex*
Locative - OBL/ADJ	*fejn* 'where'
Locative - OBL/ADJ	*mnejn* 'from where' < *minn* 'from' + *fejn* 'where'
Locative - OBL OBJ/ADJ OBJ	P + *fejn*
Internally-headed NRRCs	*liema*

Table 2: The patch-work that constitutes the employment of the *wh*-strategy in Maltese RRC/NRRCs

The above identified constraints that determine the distribution of ø-marked finite RCs could be understood as a residue of an earlier, more widely used strategy in the history of Maltese. The fact that it is constrained to indefinite antecedents is not random, since it could be a remnant of an earlier situation in Maltese when it was closer to Arabic. Indeed, a constraint still holds in different Arabic varieties to this day, whereby in the context of a an indefinite antecedent, a zero strategy is employed. Beyond this point of similarity, the rest of the constraints on the employment of this strategy in Maltese are specific to the language. In (17b) above, which I repeat below in (28) for ease of exposition, beyond the presence of a [-DEF] antecedent, we observe the requirement to have an imperfective predicate internal to the RC, namely,

joqtol ‘kill.IPFV’, as well as an antecedent which is functionally-dependent with a SUBJ in-clause function, i.e. where *tifel* ‘boy’ is not merely the SUBJ of the main clause headed by the verb *beża’* ‘fear’ but, crucially, also the in-clause SUBJ of the predicate within the RC.

(28) *Tifel* ø *j-o-qtol* *il-qtates* *mhux* *se*
boy 3M-FRM.VWL-kill.IPFV.SG DEF-cat.PL NEG PROSP
j-i-bża’ *minn* *ġurdien.*
3M-FRM.VWL-fear.IPFV.SG from mouse
‘A boy who kills cats is not going to fear a mouse.’
Borg & Azzopardi-Alexander (1997, p. 35)

If we attempt to change the RC’s predicate to one with a perfective form, as in (29a), or if we change the in-clause function to a direct object, or an object of a preposition, for instance, as in (29b-c), ungrammaticality results.

(29a) **tifel* ø *{qatel* *il-qtates}* …
boy kill.PFV.3SGM DEF-cat.PL
Intended: ‘a boy that killed cats’ *PFV predicate

(29b) **Tifel* ø *{n-af}* *qed* *j-i-studja.*
boy 1-know-IPFV.SG PROG 3M-EPENT.VWL-study.IPFV.SG
Intended: ‘A boy I know, is studying.’ *OBJ in-clause

(29c) **čavetta*$_i$ ø *{n-i-ftaħ* *il-bieb* *bi-ha*$_i$*}*
key.SGF 1-FRM.VWL-open.IPFV.SG DEF-door with-3SGF.GEN
Intended: ‘a key I open the door with’ *OBJ of P in-clause

To exemplify the whole array of the constraints that govern the employment of the zero strategy, (30) instantiates an RC introduced via this means while additionally involving a long-distance anaphoric dependency between the indefinite antecedent and a POSS in-clause function that is an argument of the OBJ *omm* ‘mother’ internal to the clausal argument embedded by the RC’s matrix (imperfective) predicate ħaseb ‘think’.

(30) Tifel$_i$ ø *{n-a-ħseb* *[li* *t-af* *lil*
boy 1-FRM.VWL-think.IPFV.SG COMP 2-know.IPFV.SG ACC
omm-u$_i$]}, *weġġa'.*
mother-3SGM.GEN be hurt.PFV.3SGM
'A boy I think you know his mother has been hurt.'
long-distance POSS in-clause function: Camilleri & Sadler (2016, p. 159)

3.3 The gap and resumptive pronoun strategies

In association with these three different strategies used to introduce RCs is the presence of either a gap, or a resumptive pronoun strategy, which this time round is present internal to the RC. (28), for instance, presented above, illustrates the presence of a gap, i.e. the absence of any overt material *in situ* at the location of the in-clause function, which happens to be the subject. The resumptive pronoun strategy, in contrast, involves the presence of a pronominal form occupying the grammatical function position internal to the clause with which the antecedent is anaphorically linked. An earlier instance of this strategy is shown in (20b), as well as (30). (31) below exemplifies the resumptive strategy in the context of all of the three RC strategies we have been looking at. (31a) illustrates the use of the *li* strategy in the context of an anaphoric dependency between the antecedent *id-dar* 'the house' and the pronominal resumptive form fulfilling the OBL OBJ function, i.e. the OBJ of the P *fi* 'in', with the PP headed by *fi* 'in' functioning as the locative OBL argument of the RC's predicate *trabba* 'bring/raise up'.

The obligatory nature of the resumptive pronoun in this *in-situ* position in Maltese follows naturally from the fact that the language does not allow P-stranding, i.e. the presence of a preposition without its associated complement *in situ*. This then explains the morphosyntactic contrast that obtains in the semantically equivalent constructions in (20) above, once the *li* strategy in (20b) substitutes the *wh*-pronoun strategy in (20a). In the former, the resumptive pronoun is obligatorily

bound to the P *ma'* 'with', while, in the latter, a gap in-clause strategy is present. (31b) is a dialectal, rather than a Standard construction, for reasons established earlier above. Nonetheless, I am providing this instance here so as to be able to display the complete paradigmatic array of contextual and structural possibilities. In this DAT-marked *wh*-pronoun introduced RRC, the antecedent is anaphorically-bound by the non- selected/ extra-argumental DAT pronoun bound onto the RC's predicate *faqa'* 'burst'.[11] (31c), on the other hand, involves the presence of a (rare) NRRC that is introduced via a zero strategy and whose indefinite antecedent is anaphorically-bound to the internal possessor function annexed in a construct state structure headed by the noun *sid* 'owner'.

(31a) *id-dar*$_i$ ***li*** *t-rabbej-t* *fi-**ha** ...*
DEF-house.SGF COMP REFL-bring up.PFV-1SG in-3SGF.GEN
'the house that I was brought up in …'
li strategy + resumptive pronoun

(31b) *ir-raġel **'il*** ***min*** *faqgħ-u-**l-u**$_i$* *l-karozza ...*
DEF-man DAT who burst.PFV.3-PL-DAT-3SGM DEF-car.SGF
Lit. 'the man to whom they burst (on-him) the car …'
wh-pronoun strategy + resumptive pronoun - (non-Standard Maltese)

(31c) *Daħl-u* *f'dar*$_i$ ø *sid-**ha**$_i$*
enter.PFV.3-PL in.house.SGF owner.SGM-3SGF.GEN
msiefer.
abroad.SGM
'They entered a house, whose owner is abroad.'
ø strategy + resumptive pronoun - Aquilina (1973, p. 338)

Constraints hold, however, as to where and when it is possible to employ a resumptive pronoun strategy. So for instance, Maltese is governed by what is in the literature referred to as the Highest Subject Restriction (Borer (1984),

11 More detail on the morphosyntax and semantics of non-selected DAT pronominal uses in Maltese can be found in Camilleri & Sadler (2012b).

McCloskey (1990)), which bars the presence of a resumptive pronoun such as *hu* 'he' in (32), in the position of the highest SUBJ within the RC.

(32)	*it-tifel*	*li*	*ø/*hu*	*ħareġ*	*issa …*
	DEF-boy	COMP	he	go out.PFV.3SGM	now
	'the boy that went out now …'				
	Highest SUBJ Restriction: gap/*resumption				

To better understand what is meant by the highest SUBJ, (32) is contrasted with (33), where this time we observe that the in-clause SUBJ function with which the antecedent displays a dependency is embedded deep within the RC; specifically as the SUBJ of the predicate ħareġ 'go out' in the embedded clause of the embedded predicate ħaseb 'think'. Such a type of dependency between the antecedent and the in-clause function is referred to as a long-distance dependency, in contrast to the immediate distance dependency that obtains vis-à-vis the in-clause SUBJ position in (32), which is in the highest (and only) clause within the RC. Since the dependency that obtains in (33) does *not* involve the highest SUBJ, the presence of a free (i.e. non-bound) resumptive pronoun in the in-clause SUBJ position becomes optionally available, and stands as a possible alternative to the gap strategy. It may well be the case that for different speakers, the resumptive pronoun strategy only becomes possible when deeper embedding is involved.

(33)	*T-kellim-t*	*ma'*	*tifel*$_i$	*{li*	*smaj-t*	*[li*
	recip-talk.IPFV-1SG	with	boy	COMP	hear.PFV-1SG	COMP
	intom	*t-af-u-(**h**)*$_i$		*sew]}*	*u*	
	you.PL	2-know.IPFV-PL-3SGM.ACC		well	CONJ	
	qal-l-i …					
	say.PFV.3SGM-DAT-1SG					
	'I talked with a boy that I heard that you (PL) know well, and he told me ...'					
	Long-distance [-DEF] OBJ: resumption/gap					

Notwithstanding the robustness of the Highest SUBJ constraint in Maltese, it can nonetheless be overridden in the context of island environments (Ross, 1967). Such environments, for our purposes here can be understood as constructions that, in Maltese and other languages that employ similar resumptive strategies, can be 'saved' via the obligatory presence of an anaphoric dependency, rather than a functional one involving a gap, and where extraction outside of them is not otherwise possible. One such instance is the Coordinated Island constraint. In such an island context, if the antecedent's in-clause function is a SUBJ, specifically an element within a set of coordinated predicates that make up the SUBJ value, i.e. *Rita u hi* in (34), the dependency involved between the antecedent, i.e. *Marija* in (34) and the in-clause grammatical function must be anaphoric, i.e. involving the obligatory presence of a resumptive pronoun, *hi* in (34), even if it happens to be in the highest SUBJ position of the RC. This is what we have in (34). The omission of the free resumptive pronoun *hi* 'she' in (34), which would have otherwise safeguarded the Highest SUBJ restriction, would have, in turn, resulted in the ungrammaticality of the whole structure.

(34) *Ma* *n-af-x* *jekk*
NEG 1-know.IPFV.SG-NEG whether
t-i-f<t>akar-x, *iżda* *Marija*$_i$ *li*
2-EPENT.VWL-remember.REFL.IPFV.SG-NEG but Marija COMP
{rita *u* ***hi***$_i$*}* *kien-u* *ħarġu* *flimkien, …*
Rita CONJ she be.PFV.3-PL go out.PFV.3-PL together,
'I don't know whether you remember, but Mary, who Rita and her had gone out together …'
Coordinate Island constraint: resumption/*gap in SUBJ

Such island environments override the general gap-resumptive pronoun distribution in other contexts. For instance, a [-DEF] OBJ in-clause function can take either a gap or a (bound) resumptive pronoun, as illustrated in (35), which specifically involves a long-distance dependency between *tifel* and (*-h*). (The same distribution holds in the immediate distance dependency counterpart.) However, in the context of what is referred to as a

Complex NP constraint, where what is involved is a RC within another RC, thus creating an even more complex NP headed by the matrix RC's antecedent, the same dependency, i.e. that between a [-DEF] antecedent and a long-distance in-clause object function, must this time round obligatorily involve a resumptive pronoun, as shown in (36).

(35) *T-kellim-t* *ma' tifel*$_i$ *{li* *smaj-t* *[li* *intom*
RECIP-talk.IPFV-1SG with boy COMP hear.PFV-1SG COMP you.PL
t-af-u-(h)$_i$ *sew]}* *u* *qal-l-i* ...
2-know.IPFV-PL-3SGM.ACC well CONJ say.PFV.3SGM-DAT-1SG
'I talked with a boy that I heard that you (PL) know well, and he told me ...'
long-distance [-DEF] OBJ: resumption/gap

(36) ... *tifel*$_i$ *{li* *smaj-t* *[li* *intom* *(huma)* *dawk*
… boy COMP hear.PFV-1SG COMP you.PL (COP.3PL) DEM.PL
{li *t-af-u-*(h)*$_i$ *sew }]}*
COMP 2-know.IPFV-PL-3SGM.ACC well
'... a boy that I heard you are those who know him well'
Complex NP Island: long-distance [-DEF] OBJ: resumption/*gap

Just as Island constraints can override the prototypical gap-resumptive pronoun distribution otherwise present in *li*-introduced RCs, the same applies in the context of RCs introduced by the *wh*-pronoun strategy. If we stick to Standard contexts (and thus remove the example in (31b) from the equation), the data in (20a), (21), and (22) all involve the presence of a gap strategy, which is indeed obligatory. The presence of island environments within the RC *changes* that distribution, such that in parallel to what we have observed in the context of *li*-introduced RCs, in the context of *wh*-pronoun introduced RCs too, an obligatory resumptive pronoun becomes necessary.

The island contexts presented this time round to illustrate this behaviour include the Adjunct Island constraint and the *Wh*-Island constraint in (37a) and (37b), respectively. The former involves a context where the in-clause function which the antecedent displays a dependency with is embedded within the ADJ-clause introduced by

qabel 'before' within the RC. The *Wh*-island context in (37b), on the other hand, involves an in-clause function that is deeply embedded within the *wh*-introduced clausal argument of the predicate *skopra* 'discover', which is itself, in turn, embedded as a clausal argument of *pprova* 'try', embedded by the RC's matrix predicate *ried* 'want'.

(37a) il-mara$_i$ *{ma' **min*** *il<t>qaj-t* *{qabel ma*
DEF-woman with who meet-RECIP.PFV-1SG before COMP
biss *kon-t* *n-af-*(**ha**)$_i$}}*
only be.PFV-1SG 1-know.IPFV.SG-3SGF.ACC
'the woman with whom I met before even knowing'
Adjunct Island constraint: resumption/*gap

(37b) *Dan* *hu* *l-post$_i$* *{**fejn*** *int*
DEM.SGM COP.3SGM DEF-place.SGM where you.SG
rid-t *darba* *[t-i-pprova*
want.pfv-2SG once 2-EPENT.VWL-try.IPFV.SG
[t-i-skopri *{jekk* *qattx* *għix-u*
2-EPENT.VWL-discover.IPFV.SG whether ever live.PFV.3-PL
fi-(**h**)$_i$* *id-dinosawr-i}]]}*
in-3SGM.GEN DEF-dinosaur-pl
'This is the place where you wanted to know whether dinosaurs ever lived in.'
***Wh*-Island constraint**: resumption/*gap

With that contained, yet comprehensive overview of the strategies employed internal to the Maltese RRCs and NRRCs, and their interaction with strategies used to introduce them, along with the constraints that govern both these types of RC strategies, we now turn our attention to the sub-types of FRCs.

4. A note on Maltese FRCs

Structurally, FRCs are special in the sense that, unlike both RRCs and NRRCs, they do not involve an identifiable antecedent, yet semantically, they behave like RRCs, rather than NRRCs, as was mentioned earlier on in §2. However, a major semantic difference which distinguishes FRCs from RRCs is the fact that plain FRCs of

the type in (3), repeated below as (38), are interpreted as definite, in line with findings in Jacobson (1995), Grosu & Landman (1998), Izvorski (2000), and Caponigro (2003), implying therefore, that a paraphrase of such FRCs is *only* possible with definite NP antecedents (39). This is in contrast with the otherwise unrestricted availability of both [+/-DEF] antecedents in the context of RRCs.

(38) *I will eat **what** they'll give me.* Plain FRC

~

*(39) I will eat **that**/*__anything__ which they'll give me.* [+DEF]-headed RRC

In English, plain FRCs contrast with *ever*-type FRCs, such as (40), which take on a distinct reading. For instance, plain FRCs are definite descriptions that can also be paraphrased by universal quantifiers. This may not necessarily be the case with *ever* type FRCs. Moreover, while plain FRCs entail or presuppose existence, this may not be the case with *ever*-type FRCs.

(40) I will eat ***whatever*** I find. ever-type FRC

While Maltese, as illustrated in Camilleri (2010), has both types of FRCs, i.e. plain ones, and *ever*-type ones, and which are even inclusive of a partially different set of *wh*-pronominal forms, so far we only have a better grasp of the semantics and (morpho) syntax of plain FRCs, as provided in Sadler & Camilleri (2018). (41a), for instance, is representative of a plain FRC in Maltese which, with its definite interpretation, can be paraphrased as in (41b). The example in (41a) illustrates how definite interpreted plain FRCs in Maltese can occur as left-dislocated topics in a construction; in this case the FRC is anaphorically-bound by the resumptive pronoun *-u* functioning as the object of the predicate *nesa* 'forget'. We will see below that this is in contrast with the inability of such a dependency in Maltese, in the case of plain FRCs interpreted indefinitely.

(41a) *T-af* *li* *[{**x'qal-l-i**}]$_i$,* *kollu*
2-know.IPFV.SG COMP what say.PFV.3SGM-DAT-1SG all
*nsej-t-**u**$_i$?*
forget.PFV-1SG-3SGM.ACC
'Do you know that I have forgotten all that he told me?'
Definite plain FRC: Borg & Azzopardi-Alexander (1997, p. 37)

~

(41b) *T-af* *li* *[dak* *{**li***
2-know.IPFV.SG COMP DEM.SGM COMP
qal-l-i}]$_i$, *kollu* *nsej-t-**u**$_i$?*
say.PFV.3SGM-DAT-1SG all forget.PFV-1SG-3SGM.ACC
'Do you know that I have forgotten all that he told me?'
Definite plain FRC: Borg & Azzopardi-Alexander (1997, p. 37)

Maltese *ever*-type FRCs, such as those of the sort represented in (11)-(12), early on in §2, and below in (42) (with (42c) functioning specifically as an adjunct *ever*-type FRC), still await a better description and analysis.

(42a) *T-i-sta'* *t-ieħu* *{**liem(a)***
2-EPENT.VWL-can.IPFV.SG 2-take.IPFV.SG whichever
t-rid}.
2-want.IPFV.SG
'You can take whichever you want.'

(42b) *I-mur* *{**fejn*** *i-mur},* *dejjem* *ħa*
3M-go.IPFV.SG where 3M-GO.IPFV.SG always PROSP
j-sib-ni *waraj-h.*
3M-find.IPFV.SG-1SG.ACC behind-3SGM.GEN
'Wherever he goes, he's always going to find me supporting him.' *ever*-type FRC

(42c) *Se* *n-a-għmel* *{**(kull) kif*** *t-għid-l-i*
PROSP 1-FRM.VWL-do.IPFV.SG however 2-say.IPFV.SG-DAT-1SG
n-a-għmel}.
1-FRM.VWL-do.IPFV.SG
'I will do however you tell me to.'[12] adjunct *ever*-type FRC

12 The use of the form *kull kif* 'however' is dialectal, and specific to the Gozitan varieties.

I will from now on concentrate entirely on plain/non-*ever* FRCs in Maltese. As established in Sadler & Camilleri (2018), this sub-set of FRCs is in Maltese *not* restricted to definite interpretations, even if the indefinite counterparts are governed by certain restrictions on their occurrence, paralleling closely (but not completely overlapping) the set of constraints presented in §3 when discussing the structural restrictions that pertain to the contexts when [-DEF] [+HUMAN] antecedents are allowed to head RCs in Maltese. Beyond this interesting fact, i.e. that two semantic readings are available to non-*ever* FRCs in Maltese, albeit governed by distinct structural conditions, the *definite* sub-set of these FRCs can in fact be introduced not solely by a *wh*-pronoun strategy (as wrongly claimed in Camilleri (2010)), but *additionally* by means of the complementiser strategy we have been observing in the context of RRCs and NRRCs in our discussion in the previous sections, i.e. by means of the complementiser *li*.

Constructions such as (43) below, which are possible in Maltese (and in fact in different Arabic varieties, too, as explicitly discussed for the first time in Sadler & Camilleri (2018)) is typologically rare, if not unique to Arabic and Maltese. The crosslinguistic literature lacks any discussion of non-*wh*-pronominal strategies for FRCs; so much so that in Caponigro's (2003) crosslinguistic study of FRCs and *wh*-items, a free relative is indeed critically defined by the occurrence of a *wh*-item. To native speakers, expositions of the set of FRCs in (43) often feel as though they lack some sort of demonstrative head, e.g *dik* 'DEM.SGF' in (43a), for example, which, once inserted, renders the whole construction into a (headed) RRC. This is one piece of syntactic proof (amongst others provided in Sadler & Camilleri (2018)) used in support of the definite semantics attributed to such complementiser introduced FRCs in Maltese (and Arabic). What is presented in (43) is an array of *li*-introduced FRCs in Maltese including ones with reference to a [+HUMAN] antecedent, as in (43a), as well as ones with a resumptive pronoun, as in (43c).

Moreover, (43a) involves a FRC that fulfills the matrix SUBJ argument, with the in-clause function being also a SUBJ; (43b) illustrates an OBJ function in both clauses; and (43c) involves a FRC that is in subject position, with the in-clause function being an object of a P.

(43a) *{**Li*** *xtra-t* *mingħand-ek},* *ġie-t*
COMP buy.PFV-3SGF from-2SG.GEN come.PFV-3SGF
s'għand-i *llum.*
until.at-1SG.GEN today
'The one who bought (something) from you came to me today.'
[+HUMAN] [+DEF] & in-clause gap: Sadler & Camilleri (2018, p. 10)

(43b) *Għamil-t* *{**li*** *għid-t-l-i}.*
do.PFV-1SG COMP say.PFV-2SG-DAT-1SG
'I did what you told me.'
[-HUMAN] [+DEF] & in-clause gap

(43c) *{Li*$_i$ *kil-na* *fi-**h**$_i$* *aħna}* *kien*
COMP eat.PFV-1PL in-3SGM.GEN we be.PFV.3SGM
vera *tajjeb.*
true good.SGM
'That which we ate in, was really good.'
[-HUMAN] [+DEF] & in-clause resumptive pronoun: Sadler & Camilleri (2018, p. 11)

Beyond the use of *li*, just as is the case in the contexts of non-FRCs as illustrated by example (18) in the introduction to the previous section, we also find the use of *milli* introduced FRCs, as is in fact documented in Sutcliffe (1936), who refers to such constructions as relatives with an 'unexpressed antecedent'. Apart from *milli* (44a), as noted by Sutcliffe himself, it is possible to additionally find the use of *għal li* (44b) in such FRC contexts, which is otherwise not an available option in the context of RRCs/NRRCs. *Għal li* this time round involves the fusion of the P *għal* 'for' and the complementiser *(il)li*. While the FRC in (44a) fulfills an OBL OBJ function as an argument of the P *barra* 'apart', the in-clause function is that of an OBJ. In (44b), the FRC fulfills the OBL function of *ġie* 'come', and the in-clause OBJ function as an argument of *xtaq* 'wish'.

(44a) *barra* *{**milli*** *għid-na}*
apart from.COMP say.PFV-1 PL
'apart from what we said'

(44b) *issa* *n-i-ġ-u* *{**għal li*** *xtaq-t}* *n-għid*
now 1-FRM.VWL-come.IPFV-PL for.what want.PFV-1SG 1-say.IPFV.SG
'now we come to what I wished to say' Sutcliffe (1936, p. 183)

Indefinite-interpreted non-*ever*-type FRCs differ from definite ones in that, while they are primarily constrained to be introduced via *wh*-pronouns, their availability in the grammar is governed by the lexical and (morpho)syntactic constraints that condition [-DEF] headed counterparts introduced by the *wh*-pronouns *x'/xi* and *(')l) min*, including the obligatory requirement for the RC's predicate to be imperfective in form. Slight differences do exist, however. As discussed in Sadler & Camilleri (2018), non-headed modal existential constructions, as plain FRCs interpreted indefinitely are referred to, can *only* function as OBJs or theme OBJs to the predicate which takes them as their argument, in contrast to the possibility of the headed counterpart to also function as that predicate's SUBJ. Moreover, while a certain lexical predicate may allow for its argument to be modified by a modal existential, that same predicate may not necessarily readily allow a non-headed modal existential construction to take the role of its own argument. Such a contrast is provided in (45), exemplified by the predicate *xtaq* 'wish'.

(45a) *N-i-x<t>ieq* *xi* *ħaġa {**x'n-a-għmel}*.
1-EPENT.VWL-wish.REFL.IPFV.SG some thing what.1-FRM.VWL-do.IPFV.SG
'I wish something to do.' Headed modal existential RC

(45b) **N-i-x<t>ieq* *{**x'n-a-għmel}*.
1-EPENT.VWL-wish.IPFV.SG what.1-FRM.VWL-do.IPFV.SG
Intended: 'I wish what to do'. *Modal existential RC: Sadler & Camilleri (2018, p. 42)

As a consequence of the constraint requiring indefinite plain FRCs to be restricted to an OBJ grammatical function of sorts,

a clitic left-dislocated construction such as that in (46) is ruled ungrammatical. This is because the FRC bears a TOPIC discourse function, rather than the OBJ grammatical function of the predicate *sab* 'find'. The predicate's object function is filled in by the bound resumptive pronoun *-u*. The ungrammaticality of this construction is in direct contrast with that in (41a), where a definite-interpreted plain FRC was shown to be able to take a TOPIC function in a clitic left-dislocated structure.

(46) **[{**X'**t-i-lbes}]$_i$* *ma*
what.2-FRM.VWL-wear.IPFV.SG NEG
*sib-t-**hu**$_i$-l-ek-x.*
encounter.PFV-1SG-3SGM.ACC-DAT-2SG-NEG
Intended: 'What/Something to wear, I didn't find-it for you.'
Indefinite plain FRC: Sadler & Camilleri (2018, p. 37)

Having highlighted some of the most salient facts about Maltese plain FRCs, I conclude this dedicated side-note on such structures, and will leave a detailed description and analysis of *ever*-type FRCs for future research.

5. Conclusion

This paper has synthesised, highlighted, rectified, sharpened, and brought together full circle in one place the main claims and findings on RCs presented in earlier works. We have seen that Maltese has (at least) three different types of RCs: RRCs, NRRCs, which can be either externally, or internally-headed, and FRCs, which in Maltese can be of the plain type, or the *ever*-type. The plain type was shown to take two distinct readings in Maltese: definite, and the more constrained, indefinite, with significant structural, semantic, and lexical constraints contrasting the latter to the former. Definite FRCs in Maltese (as in Arabic) have been shown to be quite rare typologically, in that they can be introduced

by a complementiser strategy (which includes the complementisers *li*, *milli* and *għal li*), apart from a *wh*-strategy; the latter strategy having been otherwise said to define FRCs, crosslinguistically. *Ever*-type FRCs still remain to be better described and analysed.

As we narrowed in our focus, the core of the paper elaborated upon the strategies employed to introduce RCs in Maltese, as well as those employed internal to them. We have seen that Maltese makes use of two strategies for definite FRCs: complementiser and *wh*-, three strategies for RRCs: complementiser, *wh*-, and a zero, while NRRCs rarely take a zero strategy and are otherwise introduced via the complementiser and *wh*-pronoun strategies. The latter strategy includes the *wh*-item *liema* introducing internally-headed NRRCs, and which stands in complementary distribution with the complementiser strategy in such constructions. Internal to the different RCs introduced by these distinct strategies, we have seen that either a gap or a resumptive pronoun is present in the in-clause function, i.e. the in-situ grammatical function which the antecedent is associated with internal to the RC. Stress was laid upon how the choice of these strategies, i.e. when and in relation with what other concomitant factors they occur, is highly constrained.

In having brought the different facts together in one place here, the landscape obtained allows us to make better evaluations of certain behaviours. For instance, the highly constrained (and receding) zero strategy was posited to be the result of what vestiges reside from a once fully-fledged (and systematic) functioning strategy in the system of Maltese in some earlier stages of the language, given the reflex of the zero strategy constrained to indefinite antecedents, (as is the case when it is employed in Maltese), in the rest of the Arabic system. Furthermore, the landscape obtained in this paper, based on how things currently stand in Maltese, provides us with a vantage point from where we can now characterise what prevalent gaps exist in the grammar of RCs in Maltese. A primary gap has been identified, where

it has been shown that it is essentially impossible to relativise upon a [+DEF] [+/-HUMAN] antecedent using the *wh*-pronoun strategy in Standard Maltese. The dialectal varieties, in contrast differentiate on the basis of the [+/-HUMAN] parameter, and while able to relativise [+DEF] [+HUMAN] antecedents, this is not a possibility with [-HUMAN] counterparts. Yet another feature-value based split has been singled out in the system. The negative vs. positive POLARITY values attributed to universal indefinites primarily effect their distribution as antecedents of RRC vs. NRRCs. Of most interest however is the fact that it is impossible for positive universal indefinites to be relativised upon in the first place. This is in contrast with their negative universal counterparts, which can be relativised upon strictly as antecedents of RRCs. This POLARITY-based split becomes even more stark when one observes how the reflex of this gap maps out in the system. The positive universal indefinite RRC gap is substituted by a FRC structure introduced by positive universal *ever*-type *wh*-pronouns; for which a NEG counterpart does not exist in the system. The reason(s) behind these gaps and substitutions in the system, and whether there is a connecting link beyond the feature-value [-DEF] in these two identified case, if at all semantic or (morpho)syntactic, remain(s) yet to be discovered, and understood.

Abbreviations

1, 2, 3	first, second, third person		
ACC	accusative	GEN	genitive
CAUSE	causative	IPFV	imperfective
COMP	complementizer	M	masculine
CONJ	conjunction	MASS	mass noun
COP	copula	NEG	negative
DAT	dative	PASS	passive
DEF	definite article	PFV	perfective
DEM	demonstrative	PL	plural
DU	dual	PROG	progressive
ELAT	elative	PROSP	prospective

EPENT.VWL	epenthetic vowel	RECIP	reciprocal
F	feminine	REFL	reflexive
FRM.VWL	formative vowel	SG	singular

References

Aquilina, Joseph (1973) *The structure of Maltese: A study in mixed grammar and vocabulary*. Malta: The Royal University of Malta.

Arnold, Doug (2007) Non-restrictive relatives are not orphans. *Journal of Linguistics* 43, 271–309.

Borer, Hagit (1984) Restrictive relatives in Modern Hebrew. *Natural Language and Linguistic Theory* 2, 219–260.

Borg, Albert (1991) Complementation in Maltese. In: Brincat, Joseph (ed.), *Languages of the Mediterranean: Substrata: the islands, Malta*. Malta: University of Malta, 218–227.

Borg, Albert (1994) Maltese complement sentences to non-verbs. In: Caubet, Dominique & Vanhove, Martine (eds.), *Actes des premières journées internationales de dialectologie arabe de Paris*. Paris: Publications Langues 'O, 109–120.

Borg, Albert & Azzopardi-Alexander, Marie (1997) *Maltese. Lingua descriptive grammars*. London & New York: Routledge.

Camilleri, Maris (2010) *Relative Clauses in Maltese*. University of Essex: Unpublished MRes thesis.

Camilleri, Maris (2012) The Maltese relative clause: A source of standard-dialect variation. *Ilsienna* 2, 1–12.

Camilleri, Maris (2014a) The Maltese restrictive relative clause. In: Borg, Albert, Caruana, Sandro & Vella, Alexandra (eds.), *Perspectives on Maltese linguistics*. Berlin: Akademie Verlag, 161–200.

Camilleri, Maris (2014b) *The stem in inflectional verbal paradigms in Maltese*. University of Surrey: PhD thesis.

Camilleri, Maris & Sadler, Louisa (2011) Restrictive relative clauses in Maltese. In: Butt, Miriam & King, Tracy Holloway (eds.), *Proceedings of the LFG11 Conference*. Stanford, CA: CSLI Publications, 110–130.

Camilleri, Maris & Sadler, Louisa (2012a) An LFG Approach to non-restrictive relative clauses in Maltese. University of Essex: *Research reports in linguistics.*

Camilleri, Maris & Sadler, Louisa (2012b) On the Analysis of Non-selected Datives in Maltese. In: Butt, Miriam & King, Tracy Holloway (eds.), *Proceedings of LFG12*. Stanford, CA: CSLI Publications, 118–138.

Camilleri, Maris & Sadler, Louisa (2016) Relativisation in Maltese. *Transactions of the Philological Society* 114(1), 117–145.

Caponigro, Ivano (2003) *Free Not to Ask: On the Semantics of Free Relatives and Wh-words Crosslinguistically*. University of California, Los Angeles: PhD thesis.

Cinque, Guglielmo (2008) Two types of non-restrictive relatives. In Bonami, Olivier & Hofherr, Patricia Cabredo (eds.), *Empirical Issues in Syntax and Semantics* 7, 99–138. http://www.cssp.cnrs/fr/eiss7.

Citko, Barbara (2002) (Anti) reconstruction effects in free relatives: A new argument against the Comp account. *Linguistic Inquiry* 33(3), 507–511.

Citko, Barbara (2008) An argument against assimilating appositive relatives to coordinate structures.*Linguistic Inquiry* 39(4), 633–655.

Fabri, Ray (1987) An analysis of grammatical agreement in Maltese. University of Dusseldorf: MA Dissertation.

Grosu, Alexander & Landman, Fred (1998) Strange relatives of the third kind. *Natural Language Semantics* 6, 125–70.

Huddleston, Rodney & Pullum, Geoffery K (2002) *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Izvorski, Roumyana (2000) *Free relatives and related matters*. University of Pennsylvania: PhD thesis.

Jacobson, Pauline (1995) On the quantificational force of English free relatives. In: Bach, Emmon, Jelinek, Eloise, Kratzer, Angelika, & Partee, Barbara H. (eds.), *Quantification in Natural Languages*. New York: Springer, 451–486.

Keenan, Edward L. & Comrie, Bernard (1977) Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1), 63–99.

McCloskey, James (1990) Resumptive pronouns, binding, and levels of representation in Irish. In: Hendrick, Randall (ed.), *The Syntax of the Modern Celtic Languages*, volume 23 of *Syntax and Semantics*. New York: Academic Press, 199–256.

Puech, Gilbert (1979) *Les parlers maltais: Essai de phonologie polylectale*. University of Lyon 2: PhD thesis.

Ross, John Robert (1967) *Constraints on Variables in Syntax*. MIT: PhD thesis.

Sadler, Louisa & Camilleri, Maris (2018) Free relatives in Maltese. *Brill's Annual of Afroasiatic Languages and Linguistics* 1, 1–45.

Sutcliffe, Edmund F. (1936) *A grammar of the Maltese language: With crestomathy and vocabulary*. Oxford: Oxford University Press and Humphrey Milfor.

GRAMMATICAL AGREEMENT IN MALTESE

Ray Fabri

Abstract

Maltese displays a rich scenario of phenomena related to grammatical agreement between various sources and targets involving the categories of person, number and gender within various syntactic domains. After reviewing various definitions of grammatical agreement found in the literature, this study describes and discusses local agreement phenomena in Maltese in detail, exploring both Noun Phrase internal (e.g., noun adjective agreement) and Noun Phrase external agreement (e.g., verb subject agreement), as well as long distance agreement (pronoun - antecedent) with several illustrative examples. To complete the picture, this study also looks at cases of 'quirky' agreement which includes, among others, notional or non-formal agreement.

Fil-Malti hemm għadd ta' fenomeni relatati mal-qbil grammatikali bejn elementi varji li jinvolvu l-kategoriji tal-persuna, l-għadd u l-ġens fi ħdan strutturi sintattiċi varji. Wara ħarsa fil-qosor lejn diversi definizzjonijiet tal-qbil grammatikali li nsibu fil-letteratura, dan l-istudju jiddeskrivi u jiddiskuti l-fenomeni lokali tal-qbil grammatikali fid-dettall billi jesplora l-qbil fi ħdan il-Frażi Nominali (eż., il-qbil grammatikali bejn l-aġġettiv u n-nom)

u l-qbil barra l-Fraži Nominali (eż., il-qbil grammatikali bejn il-verb u s-suġġett), kif ukoll il-qbil grammatikali mbiegħed bejn il-pronom u l-anteċedent, b'bosta eżempji. Biex l-istampa tkun kompluta, dan l-istudju jiddiskuti wkoll każijiet ta' qbil 'mhux tas-soltu', li jinkludi, fost oħrajn, qbil nozzjonali jew qbil mhux formali.

1. What is grammatical agreement?

Grammatical agreement can be said to pervade the grammar of Maltese. This chapter sets out to describe grammatical agreement in Maltese and related phenomena. Let us start by trying to understand the notion of agreement. The following are four definitions of agreement found in the literature.[1]

- Lapointe (1985: 1)
  the specific morphological form of a word appearing in a sentence *correlates* with the presence, absence, or form of some other word in the sentence.
- Corbett (2008: 4) quoting Steele (1978: 610)
  some systematic covariance between a semantic or formal property of one element and a formal property of another.
- Matasović (2018: 13)
  a [grammatical] rule that says…that the presence of the feature A (out of a limited number of features) on the lexical unit X (the *controller*) requires the presence of A on the lexical unit Y (the *target*), within a syntactic domain D.

1 See, in particular, Mel'cuk, Igor (2006) and Corbett (2008) for a detailed discussion of terms and concepts related to agreement. For grammatical agreement in Maltese, see the relevant sections in Borg & Azzopardi-Alexander (1997).

- Moravcsik (1978: 333)
  a grammatical constituent A will be said to agree with a grammatical constituent B in properties C in the language L if C is the set of meaning-related properties of A and there is a covariance relationship between C and some phonological properties of a constituent B1 across some subset of the sentences of language L, where constituent B1 is adjacent to constituent B and only meaning-related non-categorical properties of constituent B1 are the properties C.

  Chomsky (1981) characterises (rather than defines) agreement as the operation of co-indexing (through superscripts) of elements involved in certain structural relations. In Chomsky's *Minimalist Program* (see Chomsky 2000), agreement is accounted for by the *Agree* operation, where *Agree* is a relation between a probe (e.g., verb inflection) and a goal (e.g., a subject noun phrase), with the probe searching for a category (e.g. noun), which can give value to its unvalued person, number, and gender features (*phi*-features).

  Often, agreement (also called *concord*), is specifically distinguished from government (or *rection*), as the following examples show.[2]

- Robins (1971: 235)
  Concord [or agreement] may be defined as the requirement that forms of two or more words of specific word classes that stand in a specific syntactic relationship with one another shall also be characterised by the same paradigmatically marked category (or categories)…

  …government or rection may be defined as the requirement that one word of a particular class in a given syntactic construction with another word of a particular class shall exhibit the form of a specific category.

2 Note that, according to Moravscik herself, her own definition above includes both agreement and government.

- Bach (1983: 70)
  If the form of an argument depends on the properties of the functor, then we say that the functor GOVERNS the relevant feature of the argument. If the form of a functor depends on the properties of its argument expression when they are in construction, then we say that the functor *agrees* with the argument with respect to the relevant property (or feature)…

Moreover, another distinction is often made between local agreement, e.g., noun-adjective within a noun phrase (the local domain), and long-distance agreement, e.g. anaphor-antecedent (or pronominal) agreement, in which the agreeing units might be structurally distant from each other. Thus, Chomsky (1993) (and also more recent work) distinguishes between agreement and binding, with the binding relations being represented by means of subscripts and agreement relations by means of superscripts.

Clearly, two basic distinctions are therefore made in the literature, namely, (1) that between agreement (or concord) and government (or rection), on the one hand, and (2) that between agreement 'proper' and anaphoric (or pronominal) agreement, on the other.

This chapter deals with agreement phenomena in Maltese, as opposed to government (e.g. case assignment), and adopts an 'inclusive' view of agreement, i.e., it includes long distance as well as local agreement. The phenomena discussed in this chapter are based on the following characterisation of agreement by Barlow and Ferguson (1988), which succinctly captures the essence of all of the definitions given above without being too restrictive.

- Barlow & Ferguson (1988: 1)
  a grammatical element *x* matches a grammatical element *y* in property *z* within some grammatical configuration *w*

To illustrate, in (1) below, *x* can be taken to be the adjective 'tall' (*twil, twila, twal*), *y* is the noun 'child', *w* is the noun phrase (NP) made up of the head noun and the modifying adjective (e.g., *it-tifel it-twil*), and property z is number (singular/plural) and gender (masculine/feminine).

(1a) *it-tifel* *it-twil*
DEF-child.SGM DEF-tall.SGM
'the tall boy'

(1b) *it-tifl-a* *t-twil-a*
DEF-child-SGF DEF-tall-SGF
'the tall girl'

(1c) *it-tfal* *it-twal*
DEF-child.PL DEF-tall.PL
'the tall children'

The rest of this chapter is divided into two main sections, reflecting the two broad domains of agreement as suggested by Lehmann (1982), namely, NP-internal agreement, and NP-external agreement, also called 'nominal concord' and 'clausal agreement' (e.g., Sande 2019: 832 and references therein). The basis for differentiating these domains lies in the fact that, typically, NP-internal agreement generally does not involve the person category, while NP-external agreement does (with the exception of predication; more below).

2. NP-internal agreement

Within the NP in Maltese, a number of elements are involved in agreement with the head noun involving number, and gender in the singular. These are the demonstrative adjective, the modifying (attributive) adjective, the numeral and the indefinite marker *wieħed*. We discuss each of these in the following with examples.

2.1 Demonstrative adjective – noun

(2a)	*dan*	*il-ktieb*	(2b)	*dan*	*ir-raġel*
	this.SGM	DEF-book.SGM		this.SGM	DEF-man.SGM
	'this book'			'this man'	
(3a)	*din*	*is-sistem-a*	(3b)	*din*	*il-mara*
	this.SGF	DEF-system-SGF		this.SGF	DEF-woman.SGF
	'this system'			'this woman'	
(4a)	*dawn*	*il-kotba/is-sistem-i*			
	this.PL	DEF-book.PL/DEF-system-PL			
	'these books/systems'				
(4b)	*dawn*	*l-irġiel/in-nisa*			
	this.PL	DEF-man.PL/DEF-woman.PL			
	'these men/women'				

In spoken Maltese, the demonstrative adjective and the article are often blended into *dal-* [dɐl] (and corresponding assimilated article forms, e.g. *das-*) for masculine singular and for plural, and *dil-* [dɪl] (and corresponding assimilated article forms, e.g. *dis-*) for feminine singular, thus *dal-ktieb*, *dar-raġel*, *dis-sistema*, *dil-mara, dal-kotba, das-sistemi, dal-irġiel, dan-nisa*. In writing, it seems that plural *dal-* tends to be avoided. In spoken Maltese, the plural demonstrative *dawn* can also be reduced to [dɐʊ] and blended with the article as [dɐʊɪl] or [dɐʊl].

2.2 Numeral and indefinite w-ħ-d 'one/a certain'

The word *wieħed*/*waħda* can function either as the numeral one or an indefinite (perhaps better non-specific: 'a certain') marker; in the latter case, it appears pre-nominally (6), while, in the former case, it appears post-nominally in the unmarked case (5a,b) but can also occur in other positions, given the right intonation (e.g., 5c). In both cases, *w-ħ-d* agrees in gender with the noun in the singular that it occurs with, thus *wieħed* for masculine and *waħda*

for feminine. Note, moreover, that indefinite *w-ħ-d* only occurs with (governs) nouns with human referents. Interestingly this is also what happens with the *lil* case marker, which generally marks a human(ised) specific nominal object.

Numeral w-ħ-d 'one'

(5a)	*Raġel*	*wieħed*	*biss*	*ġie.*
	man.SGM	one.SGM	ONLY	come.SGM.PFV
	'Only one man came.'			
(5b)	*Mara*	*waħd-a*	*biss*	*ġie-t.*
	woman.SGF	one-SGF	only	come.SGF.PFV
	'Only one woman came.'			
(5c)	*Waħd-a*	*biss*	*ġie-t*	*mara.*
	one-SGF	only	come.SGF.PFV	woman.SGF
	'Only one woman came'			
(5d)	*Irġiel*	*ħamsa*	*ġe-w*	*biss.*
	man.PL	five	come-3PL.PFV	only
	'Only five men came'			

Indefinite w-ħ-d 'a certain one'

(6a)	*Wieħed*	*raġel*	*qal-li...*
	one.SGM	man.SGM	tell.3SGM.PFV-1SG.DO
	'A (certain) man told me...'		
(6b)	*Waħd-a*	*mara*	*qal-t-li...*
	one-SGF	woman.SGF	tell.SGF.PFV-1SG.DO
	'A (certain) woman told me...'		
(6c)	Ħames	nisa	qal-u-li...
	five	woman.PL	tell-3PL.PFV-1SG.DO
	'Five women told me...'		

In terms of distribution, the indefinite marker patterns with the set of cardinal numbers which Fabri calls C2 or intransitive numerals, such as ħamsa 'five' (5d), while, the numeral patterns with the C1 or transitive numerals, such as ħames 'five' (6c). Note

that, assuming Fabri's (1994) analysis, C2 numerals are strictly speaking not inside the NP but are free to occur anywhere within the sentence, e.g. *Irġiel ġew tnejn biss* 'Only two men came' (literally: 'men came two only' and *Raġel ġie wieħed biss* 'Only one man came' (literally: 'man came one only').

Another form of *w-ħ-d* is the plural form *uħud* 'a few/some', which combines with the preposition *minn* 'from' (cmp. English 'of') and agrees with a plural noun within the *minn* PP.

(7a)	*Uħud*	*mil-l-irġiel*	*wasl-u.*
	one.PL	from-DEF-man.PL	arrive.3PL.PFV
	'Some of the men arrived.'		
(b)	**Uħud*	*mi-r-raġel*	*wasal.*
	one.PL	from-DEF-man.SGM	arrive.3SG.PFV

The rest of the numerals do not show any gender variation in form but, in terms of number, the C1 (transitive) numerals from żewġ (or ġiex) 'two' to *għaxra* 'ten' occur with a plural noun (6c), while the rest, i.e., from ħdax 'eleven' onwards, occur with a singular noun (8). If the number is over a hundred and ends in a digit from two to ten, then the noun is plural (8c).

(8a)	*ħdax-il*	*raġel*	(8b)	*mitt* *mara*
	eleven	man.SGM		hundred woman.SGF
	'eleven men'			'a hundred women'

(8c)	*mija*	*u*	*tliet*	*irġiel*
	hundred	and	three	man.PL
	'one hundred and three men'			

It seems correct to conclude that, in this case, while gender is a matter of agreement (a symmetrical relation), number is a matter of government (an asymmetrical relation), i.e. the numeral selects the form of the noun it governs, rather than co-varies with it. Note, however, that one can argue that *w-ħ-d* as an indefinite marker is different from the other transitive numerals (żewġ, *tliet*,

etc.) because, in fact, it is not a numeral and, therefore, cannot be assumed to govern its noun in terms of number rather than agree with it. This issue remains open for now.

2.3 Noun and adjective

Generally, the adjective in Maltese is post-nominal and it agrees with the noun in gender and number, as can be seen in (9).

(9a) *it-tifel* *it-twil*
DEF-child.SGM DEF-tall.SGM
'the tall boy'

(9b) *it-tifl-a* *t-twil-a*
DEF-child-SGF DEF-tall-SGF
'the tall girl'

(9c) *it-tfal* *it-twal*
DEF-child.PL DEF-tall.PL
'the tall children'

(9d) *il-karozz-a* *l-ħamr-a* *l-modern-a*
DEF-car-SGF DEF-red-SGF DEF-modern-SGF
'the modern red car'

This is also the case when the adjective occurs post-nominally in a comparative or superlative construction together with the prenominal comparative marker *aktar* (or *iżjed*) 'more/most' (10).

(10a) *l-aktar* *ktieb* *ġdid*
DEF-more book.SGM new.SGM
'the newest book'

(10b) *l-aktar* *storj-a* *ġdid-a*
DEF-more story-SGF new-SGF
'the newest story'

A number of gradable adjectives are inflected for comparison, in which case they do not require *aktar* and they are found pre-

nominally. Such adjectives do not change form for gender and number, and so do not display agreement. Note that not all gradable adjectives have a comparative form, e.g., *għajjien* 'tired', *marid* 'sick', *antik* 'old', *intelliġenti* 'intelligent', among others (cmp. English 'nicer' vs. 'more interesting').

(11) *l-isbaħ* *ktieb/* *storj-a/* *kotba*
DEF-nice.CMP book.SGM/ story-SGF/ book.PL
'the nicest book, story, books'

Therefore, a number of adjectives display four different word forms, e.g., *sabiħ* 'nice.SGM', *sabiħ-a* 'nice-SGF', *sbieħ* 'nice.PL' and *isbaħ* 'nice.CMP', while others display three forms, namely, *marid* 'sick.SGM', *marid-a* 'sick-SGF'and *morda* 'sick.PL'. Moreover, another set of adjectives, such as *interessanti* 'interesting' and *speċjali* 'special', only occur in one form and, therefore, are not specified for either gender or number, or comparison.[3]

(12) *ktieb/* *storj-a/* *kotba* *interessanti/* *speċjali*
book.SGM/ story-SGF/ book.PL interesting/ special
'interesting/special book, story, books'

Note that a small number of adjectives that do occur pre-nominally, such as *povru* 'poor', *uniku* 'unique' and *allegat* 'alleged', also generally agree with the noun in gender and number, which means that agreement, at least in these cases, is not tied to a specific syntactic position (pre- or post-nominal). However, interestingly, in pre-nominal position agreement in gender can be suspended sometimes, for example, with masculine *l-uniku* also occurring with a feminine noun (3d).[4] This is not possible with every pre-nominal adjective, so that *povru/a*,

3 These are mostly adjectives of Romance origin ending in *-i*. Note also the ċuċ 'stupid', which has a plural form ċwieċ but no 'standard' feminine form ċuċa, although one can occasionally hear this form being used.

4 See Amaira and Borg (2020) for a discussion of such 'mismatches'.

for example, has to agree in gender with the noun it modifies. Finally, it is also possible to have an adjective like *l-unika* occur with a plural noun instead of *l-uniċi* (3e); this, however, is also the case with a number of post-nominal adjectives discussed below (see (14)).

(13a) *l-unik-u* *ktieb*
DEF-only-SGM book.SGM
'the only book'

(13b) *l-unik-a* *ittr-a*
DEF-only-SGF letter-SGF
'the only letter'

(13c) *l-uniċ-i* *stejjer*
DEF-only-PL story.PL
'the only stories'

(13d) *l-unik-u* *ħaġ-a*
DEF-only-SGM thing-SGF
'the only thing'

A number of feminine singular adjectives which end in *-a* and have a plural in *-in*, including passive participles (*magħmul* 'made.SGM', *magħmul-a* 'made-SGF', *magħmul-in* 'made-PL'), can also co-occur with a plural noun in free variation with the plural *-in* form. Thus, for example, the singular feminine form *għajjiena* 'tired' can occur with a plural noun (14a,b), while the same is not possible with the adjective *xiħ* 'old' (14c).

(14a) *nisa* *għajjien-a*[5]/ *għajjen-in*
women.PL tired-PL/ tired-PL
'tired women'

(14b) *irġiel* *għajjien-a/* *għajjen-in*
man.PL tired-PL/ tired-PL
'tired men'

(14c) *nisa* **xiħ-a/* *xjuħ*
woman.PL old-SGF/ old-PL
'old women'

(14d) *irġiel* **xiħ-a/* *xjuħ*
man.PL old-SGF/ old-PL
'old men'

5 The *-a* is glossed as SGF in the relevant context (e.g., *mara għajjien-a* 'woman tired-SGF').

Finally, a marked construction must also be mentioned in which certain adjectives which normally occur post-nominally occur before a possessive noun, such as *missierek* 'your father', *ommok* 'your mother' or *wiċċek* 'your face', with the resulting construction generally carrying a negative (often ironic) meaning, as in (15).

(15a) *dik* *il-ħelw-a* *omm-ok*
that.SGF DEF-sweet-SGF mother-2SG.POSS
'that sweet mother of yours'

(15b) *dak* *il-ħelu* *missier-ek*
that.SGM DEF-sweet.SGM father-2SG.POSS
'that sweet father of yours'

(15c) *dak* *is-sabiħ* *wiċċ-ek*
that.SGM DEF-beautiful.SGM face-2SG.POSS
'that beautiful face of yours'

Note that the final *-a* of the feminine in these constructions tends to be elided when spoken, thus *dik il-ħelw'ommok*.

Maltese also has a set of collective (uncountable or mass) nouns which form part of a tripartite inflectional system consisting of (1) a singulative form which is feminine singular for agreement purposes (16b), (2) a plural form which triggers plural agreement (16c), and (3) a collective form which is masculine singular in terms of agreement (16a).

(16a) *it-tadam* *sabiħ*
DEF-tomato.SGM.COLL beautiful.SGM
'the beautiful tomato/es'

(16b) *it-tadam-a* *sabiħ-a*
DEF-tomato-SGF beautiful-SGF
'the beautiful tomato'

(16c) *it-tliet* *tadam-iet* *sbieħ*
DEF-three tomato-PL beautiful.PL
'the beautiful tomatoes'

Collective nouns trigger singular masculine agreement on the adjective as well as on any other agreeing elements, such as demonstratives within the NP and verbs NP-externally (see below for verb agreement).

To finish this section, it is worth briefly mentioning the definite article, which, as in Modern Standard Arabic (MSA), can appear on both the noun and the adjective. However, unlike MSA, in which the adjective must 'agree' with the noun in definiteness (17), in Maltese, an indefinite adjective can co-occur with a definite noun (18), and therefore definiteness must be excluded from agreement in Maltese.

(17a) *qamar-u-n* *kabiir-u-n*
moon-NOM-INDEF big-NOM-INDEF
'a big moon'

(17b) *'al-qamar-u* *'al-kabiir-u*
DEF-moon-NOM DEF-big-NOM
'the big moon'

(17c) **'al-qamar-u* *kabiir-u-n*
moon-NOM-INDEF DEF-big-NOM

(17d) **qamar-u-n* *'al-kabiir-u*
moon-NOM-INDEF DEF-nom-NOM

(18a) *il-ktieb* *il-ġdid*
DEF-book DEF-new
'the new book'

(18b) *il-ktieb* *ġdid*
DEF-book new
'the new book'

(18c) *ktieb* *ġdid*
DEF-book new
'a new book'

(18d) **ktieb* *il-ġdid*
book DEF-new

The factors that account for the asymmetric cases in (18) are discussed in Fabri (1993, 2001).

3. NP-external agreement

The agreement relation of elements outside of the NP generally involves person (first, second, third) as well as gender and number. We start off with subject verb agreement.

3.1 Subject verb agreement

In Maltese, the verb always agrees with the subject in person (1st, 2nd, 3rd), number (sg, pl) and gender (masculine, feminine) in the 3rd person singular. The following (19) is the verb paradigm for the imperfect and perfect of the so-called sound verb[6] *ħaseb* 'think', followed by some example sentences as illustration of verb subject agreement (20).

(19) Paradigm of the imperfect and perfect of the sound verb *ħaseb* 'think'.

	ħsb 'think'			
	PERFECT		IMPERFECT	
	SINGULAR	PLURAL	SINGULAR	PLURAL
1	naħseb	naħsbu	ħsibna	ħsibt
2	taħseb	taħsbu	ħsibt	ħsibtu
3M	jaħseb	jaħsbu	*ħaseb*	ħasbu
3F	taħseb		ħasbet	

(20a) *Jien* *n-aħseb* *li* *Marija* *ħarġ-et.*
I 1SG.IPVF-think that Mary go.out-3SGF.PFV
'I think that Mary went out.'

(20b) *Intom* *t-af-u* *li* *t-tfal* *marr-u* *l-Belt?*
You.PL 2.IPFV-know-PL that DEF-child.PL go-3PL.PFV DEF-city
'Do you know that the children went to Valletta?'

6 Traditional grammar distinguishes between sound and weak (defective) verbs, with the latter displaying a weak consonant (semi-vowel) *j* or *w* as one of their root consonants (e.g. *wasal* 'arrive', *bies* 'kiss').

In terms of morphology, the verb in Maltese displays a richness of forms resulting from inflectional classes (e.g., sound/weak verbs), together with morpho-phonological effects (e.g., assimilation, syllabification) involving both stems and affixes. The following table shows the allomorphy in the affixes which are involved in subject verb agreement.

(21) Subject agreement affixes: allomorphy

1 (i)n/m/r/l-	1/2SG -(Vj)t
2 (i)t, s, x, z, ż, ġ, ċ, d[7]	3F -Vt
3M j/i-	1PL -(Vj)na
3F (i)t, s, x, z, ż, ġ, ċ, d	2PL -(Vj)tu
PL -u/-Vw	

Like other Semitic languages, Maltese does not have a specific morphological infinitive verbal form. Every finite verb (i.e., not participles) is always marked for person, number, and gender. As a result, verb sequences are formed which contain verbs agreeing with the subject and, therefore, with each other[8], as can be seen in the following.

(22a) *t-rid*
3SGF.IPFV-want
'she wants'

(22b) *Marija* *t-rid* *t-oħroġ.*
Mary 3SGF.IPFV-want 3SGF.IPFV-go.out
'Mary wants to go out.'

(22c) *Marija* *t-rid* *t-oħroġ* *t-ixtr-i.*
Mary 3SGF.IPFV-want 3SGF.IPFV-go 3F-buy-SG
'Mary wants to go out to buy/to go shopping.'

7 The following are the IPA symbols corresponding to the grapheme [n, m, r, t, s, ʃ, ts, dʒ, tʃ, d, y]. V stands for 'vowel'.

8 See Maas (2009), Stolz (2009), Fabri & Borg (2017) and Azzopardi (2019) for detailed discussions of this construction.

(22d) **Marija* *t-rid* *j-oħroġ.*[9]
Mary 3SGF.IPFV-want 3SGM.IPFV-go

Note that, unlike Modern Standard Arabic,[10] in Maltese agreement of the verb with the subject in terms of person, number and gender is obligatory, no matter what the word (constituent) order is, or how many elements intervene in between.

(23a) *Marija* *marr-et.*
Mary go-3SGF.PFV
'Mary went.'

(23b) *Marr-et* *Marija.*
go-3SGF.PFV Mary
'Mary went.'

(23c) *Il-mara* *li* *ltqaj-t* *magħ-ha* *lbieraħ*
DEF-woman.SGF that meet-1SG.PFV with-3SGF yesterday
filgħaxija, *kmieni* *dalgħodu* *siefr-et.*
evening early this morning go.abroad-3SGF.PFV
'The woman I met yesterday evening went abroad early this morning.'

The verb in a relative clause relativising the subject also agrees in number, gender and person with the subject.[11]

(24a) *il-mara* *li* *feṫ-et* *il-bieb*
DEF-woman.SGF that open-3SGF.PFV DEF-door
'the woman who opened the door'

(24b) *ir-raġel* *li* *fetaħ* *il-bieb*
DEF-man.SGM that open.3SGM.PFV DEF-door
'the man who opened the door'

9 As opposed to *Marija trid li j-oħroġ* 'Mary wants that he goes out' and *Marija t-ridu j-oħroġ* 'Mary wants him to go out', which are acceptable.

10 In MSA, agreement with the subject differs, depending on word order: SVO involves agreement in person, number and gender, while VSO involves only person and gender.

11 See and Camilleri (2014), and Camilleri and Sadler (2011) and (2016) on relative clauses in Maltese.

In so-called Exceptional Case Marking (ECM) constructions with object control (25), the verb in an embedded clause, i.e. the second verb (here *jilgħab)*, agrees with its (understood) subject, which appears as the direct object in the matrix clause (in (25) *jien rajt lil Pawlu)***.**

(25) ECM: object control

Jien	*raj-t*	*lil*	*Pawlu*	*j-ilgħab.*
I	saw-1SG.PFV	CS	Paul.3SGM	3SGM.IPFV-play

'I saw Paul play/ing.'

3.2 Object verb agreement

The Maltese verb can also agree with both the direct and indirect object through pronominal clitics, traditionally known as *il-pronomi mehmużin* 'bound pronouns'.[12] The following lists the clitics attached to the verb and their allomorphs. The indirect object clitics are the same as the direct object clitics but preceded by *-l-,* which is related to the case maker *lil*, which marks a specific, human direct object (theme, patient) NP and an indirect object (recipient/benefactive) NP.

(26) The verb clitics

		DIRECT OBJECT	INDIRECT OBJECT
SINGULAR	1	-ni	-li
	2	-k/-Vk	-lVk
	3M	-u/h/hu	-lu
	3F	-ha/hie/hi	-lha/lhie
PLURAL	1	-na/nie	-lna/lnie
	2	-kom	-lkom
	3	-hom	-lhom

12 See Camilleri (2009) for a study on clitics in Maltese, and Fabri (1993) and Camilleri (2011) on pronominal clitics.

One of the functions of the pronominal clitic on the verb is to mark a topic, i.e., given or familiar information, and is typically used as shown in (27), which provides a topic context for the direct object *il-ktieb* 'the book', as opposed to (28), which provides a focus context.[13]

(27a) *U* *dak* *il-ktieb* *li* *xtra-jt* *ilbieraħ?*
and that.SGM DEF-book that buy-3SGM.PFV yesterday
'And what about that book you bought yesterday?'

(27b) *Toni* *diġà* *qra-h* *il-ktieb.*
Tony already read.3SGM.PFV-3SGM.DO DEF-book
'Tony has already read the book.' (The book, Tony already read it.)

(27c) *#Toni* *qara* *l-ktieb.*
Tony read.3SGM.PFV DEF-book
'Tony read the book.'

(28a) *X'* *qara* *Toni?*
What read.3SGM.PFV Tony
What did Tony read?

(28b) *Toni* *qara* *l-ktieb.*
Tony read.3SGM.PFV DEF-book
'Tony read the book.'

(28c) *#Toni* *qra-h* *il-ktieb.*
Tony read.3SGM.PFV-3SGM.DO DEF-book
'Tony read the book.' ('Tony read it, the book.')

As can be seen from the example above, the clitic agrees in person (1, 2, 3), number (sg, pl) and gender (m, f) with the object NP. The following are examples with topic indirect object (29b), and with both topic direct and indirect object (29c). Again, agreement guarantees that the topic NP is identified as such.

(29a) *Jien* *bgħat-t* *il-ktieb* *lil* *Pietru.*
I send-1SG.PFV DEF-book.SGM CS Peter
'I sent the book to Peter.'

13 The symbol '♯'stands for infelicitous (unacceptable in the discourse context) as opposed to ungrammatical (unacceptable in any context, marked by '*').

(29b) *Jien* *bgħat-t-lu* *l-ktieb* *lil* *Pietru.*
I send-1SG.PFV-3SGM.IO DEF-book.SGM CS Peter
'I sent the book to Peter.' ('Peter, I sent him the book')

(29c) *Jien bgħat-t-hu-lu* *l-ktieb* *lil* *Pietru.*
I send-1SG.PFV-3SGM.DO-3SGM.IO DEF-book.SGM CS Peter
'I sent the book to Peter.' ('Peter, I sent him the book')

Once a clitic is attached to the verb, word order is totally free without the need for a marked intonation.[14] Note also that the clitic has full pronominal status and, as such, can occur without the explicit object NP, as can be seen in (30).

(30a) *Toni* *qra-h.*
Tony read.3SGM.PFV-3SGM.DO
'Tony read it.'

(30b) *Jien* *bgħat-t-hu-lu.*
I send-1SG.PFV-3SGM.DO-3SGM.IO
'I sent it to him.'

Clitics can also be attached to prepositions (31) and nouns (32), thus triggering agreement with the object of the preposition and the possessor NP of the possessive noun phrase in construct. Apart from first person singular *-i*, the clitics on nouns and prepositions are the same as those attached to the verb as direct object (26), which is *-ni*, thus, e.g., *seraq-ni* 'he robbed me' but *fuq-i* 'on me'.

(31a) *Klara* *qabż-et* *fuq* *il-ħalliel.*
Klara jump-3SGF.PFV on DEF-thief.SGM
'Klara jumped on the thief.'

(31b) *Klara* *qabż-et* *fuq-u* *l-ħalliel.*
Klara jump.3SGF.PFV on-3SGM.PRPO DEF-thief.SGM
Klara jumped on the thief.' ('Klara jumped on him, the thief')

14 See Fabri (1993), Fabri and Borg (2002), Čeplö (2018) for detailed discussions of word order. See Fabri (1993) and Čeplö (2014) for studies on so-called 'clitic doubling' or 'object reduplication' like those discussed here.

(32a) *Xagħar* *it-tifl-a* *twil.*
hair.SGM DEF-child-SGF long.SGM
'The girl's hair is long.'

(32b) *Xagħar-ha* *t-tifl-a* *twil.*
hair.SGM-3SGF.POSS DEF-child-SGF long.SGM
'The girl's hair is long.' ('Her hair is long, the girl.')

Note that, as in the case of verbs, once the clitic is attached to the preposition and to the possessed noun, the complement NP, i.e. the possessor NP and the object of the preposition, is not required to be strictly adjacent to the preposition or possessed noun but can appear anywhere within the sentence (33).

(33a) *Il-ħalliel* *Klara* *qabż-et* *fuq-u.*
DEF-thief Klara jump.3SGF.PFV on-3SGM.PRPO
'Klara jumped on the thief.' ('The thief, Klara jumped on him')

(33b) *Klara* *l-ħalliel* *qabż-et* *fuq-u.*
Klara DEF-thief.SGM jump-3SGF.PFV on-3SGM.PRPO
'Klara jumped on the thief.' ('Klara, the thief, she jumped on him')

(34a) *It-tifl-a* *xagħar-ha* *twil.*
DEF-child-SGF hair.SGM-3SGF.POSS long.SGM
'The girl's hair is long.' ('The girl, her hair is long.')

(34b) *Xagħar-ha* *twil* *it-tifl-a.*
hair.SGM-3SGF.POSS long.SGM DEF-child-SGF
'The girl's hair is long.' ('Her hair is long, the girl.')

3.3 Primary & secondary predication

3.3.1 *Primary predication*

Primary predication includes sentences with a copula or other predicating verbs, such as *baqa'* 'remain' and *sar* 'become', which can have an adjective (AP) or a noun (NP) in predicate position.

Copula constructions can be either predicative (35), or equative (36). In every case, the adjective or noun in predicate position must agree with the subject in gender and number.

(35a) *L-idea* *kien-et* *tajb-a.*
DEF-idea.SGF be-SGF.PFV good-SGF
'The idea was good.'

(35b) *Il-ktieb* *kien* *tajjeb.*
DEF-book.SGM be.SGM.PFV good.SGM
'The book was good.'

(35c) Ħi-ja kien tabib.
brother.SGM-1SG.POSS be-SGM.PFV doctor.SGM
'My brother was a doctor.'

(35d) *Oħt-i* *kien-et* *tabib-a.*
sister.SGF-1SG.POSS be-3SGF.PFV doctor-SGF
'My sister was a doctor.'

(36a) *It-tabib* *kien* *ħi-ja.*
DEF-doctor.SGM be.3SGM.PFV brother.SGM-1SG.POSS
'The doctor was my brother.'

(36b) *It-tabib-a* *kien-et* *oħt-i.*
DEF-doctor-SGF be-3SGF.PFV sister.SGF-1SG.POSS
'The doctor was my sister.'

The following are examples with *baqa'* 'remain'.

(37a) *Il-kejk* *baqa'* *tajjeb.*
DEF-cake.SGM remain.3SGM.PFV good.SGM
'The cake remained good.'

(37b) *Il-pudin-a* *baqgħ-at* *tajb-a.*
DEF-cake-SGF remain.3SGF.PRV good.SGF
'The pudding remained good.'

Note that predication does not include agreement in person since, e.g., the subject can be any person (first, second, third),

while the predicative noun is generally 3rd person for the purposes of agreement, thus, e.g., *jien kont tabib* 'I was a doctor'. In other words, although external to the NP, predicate agreement functions like NP-internal agreement in not involving the person category. Note that a noun (NP) is considered 3rd person because it triggers 3rd person agreement on the verb.

(38a) *It-tabib* *ċempel.*
DEF-doctor.SGM phone.3SGM.PFV
'The doctor phoned you.'

(38b) **It-tabib* *ċempil-t.*
DEF-doctor.SGM phone-1/2SG.PFV

3.3.2 *Secondary predication*

Agreement in gender and number also takes place in secondary predication, a construction in which an adjective is predicated of the subject or object but is not the primary predicate of the clause. It can obtain a resultative or depictive interpretation. For example, in (39a) the primary predicate is *kilt* 'I ate'; the secondary predicate is *kiesaħ* 'cold' and it is predicated of the object *soppa* 'soup', obtaining a depictive interpretation. In contrast, (40) obtains a resultative interpretation. The adjective *għarwien* 'naked' in (41) is predicated of the subject NP, and is depictive.

(39a) *Jien* *kil-t* *is-sopp-a* *kiesħ-a.*
I ate-1SG.PFV DEF-soup-SGF cold-SGF
'I ate the soup cold.'

(39b) *Int* *kil-t* *l-għaġin* *kiesaħ.*
you ate-2SG.PFV DEF-pasta.SGM cold.SGM
'You ate the pasta cold.'

(40a) *Harry* *żeba'* *l-kamr-a* *safr-a.*
Harry painted.3SGM.PFV DEF-room-SGF yellow-SGF
'Harry painted the room yellow.'

(40b) *Harry* *żeba'* *l-ħajt* *isfar.*
Harry painted.3SGM.PFV DEF-wall.SGM yellow.SGM
'Harry painted the wall yellow.'

(41a) *Pawlu* *żifen* *għarwien.*
Paul dance.3SGM.PFV naked.SGM
'Paul danced naked.'

(41b) *Moira* *żifn-et* *għarwien-a.*
Moira dance.3SGF.PFV naked-SGF
'Moira danced naked.'

Another interesting case is the word *waħdu* 'alone/on his own', which has the same distribution as a secondary predicate, but which obligatorily occurs with an object clitic and, therefore, encodes person as well as number and gender information. Note that adjectives and adverbs are not inflected for person, and yet it seems that *waħdu* functions as an adverb or adjective in these constructions. In any case, it always has to agree in person, number and gender with the NP it modifies.

(42a) *Pawlu* *żifen* *waħd-u.*
Paul dance.3SGM.PFV alone-3SGM
'Paul danced on his own/alone.'

(42b) *Int* *żfin-t* *waħd-ek.*
you dance.2SG.PFV alone-2SG
'You danced on your own/alone.'

(43a) *Jien* *ra-jt* *lil* *Pawlu* *waħd-u.*
I see-1SG.PFV CS Paul alone-3SGM
'I saw Paul on his own/alone.'

(43b) *Int* *ra-jt-ni* *waħd-i.*
you see-2SG.PFV-1SG.DO alone-1SG
'You saw me on my own/alone.'

(44a) *Intom* *kil-tu* *l-kejk* *waħed-kom.*
you.PL eat-2PL.PFV DEF-cake alone-2PL
'You(pl) ate the cake on your own/alone.'

(44b) *Int* *kil-t* *il-kejk* *waħd-ek.*
you eat-2SG.PFV DEF-cake alone-2SG
'You(sg) ate the cake on your own/alone.'

w-ħ-d can also occur in constructions like (45), where, presumably, it modifies the noun within the NP.

(45) Ġanni waħd-u ma j-ista' j-agħmel xejn.
John alone-3SGM NEG 3SGM.IPFV-can 3SGM.IPFV-make nothing
'On his own, John cannot do anything.'

Other elements that resemble *w-ħ-d* in terms of their agreement patterns are the quantifiers *koll-* 'all' and *nofs* 'half', which also agree with the NP they quantify through pronominal clitics.

(46a) *Il-logħb-a* *koll-ha/nofs-ha* *kien-et* *tajb-a.*
DEF-game-SGF all-3SGF/half-3SGF be.pst-3SGF good-SGF
'The whole game was good.'

(46b) *Il-film* *koll-u/nofs-u* *kien* *tajjeb.*
DEF-film.SGM all-3SGM/half-3SGM be.PST-3SGM good.SGM
'The whole film was good.'

(46c) *Il-logħb-iet* *koll-ha* *kien-u* *tajb-in.*
DEF-game-PL all-3PL be.PST-3PL good-PL
'All the games are good.'

(46d) *Intom/aħna* *koll-ha* *tajb-in.*
you.PL/we all-3PL good-PL
'All of you/us are good.'

There are a number of anomalies in these cases. The agreement marker is not the one typical of adjectives (consonant for masculine, *-a* for feminine, *-in* or *-a* for plural) but the object clitic, as is suggested by the orthography but also by the fact that masculine singular is *-u*, which is never the case with adjectives, although it can be a masculine marker for nouns (e.g., *ziju* 'uncle'). What appears to be the third person feminine marker as a clitic (*-ha*)

here marks plural for first (*aħna*), second (*intom*) and third person, as well as for feminine singular.

One will also need to distinguish between the use of *koll-* and *nofs* as quantifiers, on the one hand, and as pseudo-predicates (see 3.5.1 below), on the other. The latter agree also in person, as in *Int kollok problemi* 'You are all (full of) problems'. However, this distinction requires careful study and I will not go into any further here.

3.4 Pronoun – antecedent

Maltese is a pro-drop language, i.e., generally does not overtly express (therefore 'drops') the pronominal subject, which is, therefore, 'understood' through the gender, number and person features on the verb (e.g., ħareġ 'he went out' in the second sentence in (47a)). The verb, therefore, agrees with the antecedent of the unexpressed pronoun (the subject of the first sentence Ġanni), which can be located anywhere within the discourse context outside of the sentence or clause containing the relevant verb.

(47a)	Ġanni	m-hu-x	hawn.	Ħareġ	j-ixtri.
	John	NEG-be.3SGM-NEG	here	go.out.3SGM.PFV	3SGM.IPFV-buy
	'John is not here. He went shopping.'				

(47b)	*Iltqaj-t*	*ma'*	*Marija.*	*Ħarġ-et*	*t-ixtri.*
	meet-1SG	with	Mary	go.out-3SGF.PFV	3SGF.IPFV-buy
	'I met Mary. She went shopping.'				

(47c)	*Mor-na*	*l-Belt.*
	go-1PL.PFV	DEF-city
	'We went to Valletta.'	

The same applies to objects, with the pronominal function borne by the object clitics (*-u* in (48a) and *-hie* in (48b)).

(48a) Ġanni m-hu-x hawn?
John NEG-be.3SGM-NEG here
Le, ma ra-jt-u-x.
no NEG see-1SG.PFV-3SGM.DO-NEG
'Isn't John here? No, I haven't seen him.'

(48b) *Marija m-hi-x hawn?*
Mary NEG-be.3SGF-NEG here
Le, ma raj-t-hie-x.
no NEG see-1SG.PFV-3SGF.DO-NEG
'Isn't Mary here? No, I haven't seen her.'

The verb can also agree with any of a set of emphatic pronouns shown in (49) and exemplified in (50).

(49) emphatic pronouns

SUBJECT	**OBJECT (direct and indirect)**
jien/a 'I'	lili 'me'
int/i 'you'(sg)	lilek 'you(sg)'
hu/wa 'he'	lilu 'him'
hi/ja 'she'	liha 'her'
aħna 'we'	lilna 'us'
intom 'you(pl)'	lilkom 'you(pl)'
huma 'they'	lilhom 'them'

(50) *Jien ma ra-jt-x lilu; ra-jt lilha.*
I NEG see-1SG.PFV him see-1SG.PFV her
'I didn't see him; I saw her.'

Here we should also mention the preposition *ta'* 'of', which, added to a clitic, is interpreted as possessive adjective (51) and possessive pronoun (52). The following is the relevant paradigm.

(51a) *Pawlu biegħ id-dar tiegħ-u.*
Paul sell.3SGM.PFV DEF-house of-3SGM
'Paul sold his house.'

(51b) *Marija* *biegħ-et* *id-dar* *tagħ-ha.*
Mary sell-3SGF.PFV DEF-house of-3SGF
'Mary sold her house.'

(52a) *Tiegħ-i* *isbaħ* *minn* *tiegħ-ek.*
of-1SG nice.CMP from of-2SG
'Mine are/is nicer than yours.'

(52b) *Tagħ-na* *għad-hom* *ma* *wasl-u-x.*
of-1PL still.3PL NEG arrive-3PL.PFV-NEG
'Ours haven't arrived yet.'

The following is the paradigm for the possessive pronoun/adjective *ta'*+clitic

(53) Paradigm for *ta'* 'of'[15]

SINGULAR	1	tiegħ-i 'my/mine'
	2	tiegħ-ek 'your/s'
	3M	tiegħ-u 'him/his'
	3F	tagħ-ha 'her/s'
PLURAL	1	tagħ-na 'our/s'
	2	tagħ-kom 'your/s'
	3	tagħ-hom 'their/s'

Finally, reflexivity in Maltese can be expressed in three ways: (1) through the combination of *lil* with a pronominal clitic, together with *nifs* 'breath' with an attached pronominal clitic (54), (2) through the combination of *ruħ* 'soul' with a pronominal clitic (55), and (3) by means of the derivational prefixes *n-* (the 7th form; see (56)), *t-* (the 5th form), and at least one case of the 6th form (t+3rd form).[16] In types (2) and (3) reflexives agree through

15 This means that, e.g., *tiegħi* 'of-me, my/mine' is a preposition or prepositional phrase which functions as a possessive adjective/pronoun. Compare to other [preposition + clitic] combinations, such as *ma'+i = miegħi* 'with-me', *magħ-kom* 'with-you(pl)'.

16 I will not go into a discussion of the relation between reflexive, passive, and middle constructions, which belongs to the sphere of meaning (semantics)

the pronominal clitic with their antecedents in gender, number and person within the clause while type (3) involves 'normal' subject-verb agreement. The following are examples.

(54a) *Jien* *ra-jt* *lil-i* *nnifis-i.*
I see-1SG.PFV CS-1SG breath-1SG
'I saw myself'

(54b) *Hi* *ra-t* *lil-ha* *nnifis-ha.*
I see-3SGF.PFV CS-3SGF breath-3SGF
'She saw herself'

(54c) *Huma* *ra-w* *lil-hom* *infus-hom.*
they see-3PL.PFV CS-3PL breath-3PL
'They saw themselves'

(55a) *Jien* *sib-t* *ruħ-i* *m-aqful* *ġo* *kamra.*
I find-1SG.PFV soul-1SG PART-lock.SGM inside room
'I found myself locked up in a room.'

(55b) *Hi* *sab-et* *ruħ-ha* *m-aqful-a* *ġo* *kamra.*
She find-3SGF.PFV soul-3SGF PART-lock-SGF inside room
'She found herself locked up in a room.'

(55c) *Huma* *sab-u* *ruħ-hom* *m-aqful-in* *ġo* *kamra.*
they find-3PL.PFV soul-3PL PART-lock.PL inside room
'They found themselves locked up in a room.'

(56a) *Jien* *n-in-ħasel* *kuljum.*
I 1SG.IPFV-REFL-wash every day
'I wash myself/get washed every day.'

(56b) *Hi* *t-in-ħasel* *kuljum.*
she 3SGF.IPFV-REFL-wash every day
'She washes herself/gets washed every day.'

(56c) *Huma* *j-in-ħasl-u* *kuljum.*
they 3.IPFV-REFL-wash-PL every day
'They wash themselves every day.'

and not morpho-syntax strictly speaking. See Spagnol (2011) for a detailed discussion.

3.5 Quirky agreement

In this section, we look briefly at a number of cases of agreement which, for various reasons, do not follow the general pattern described up to now, starting with pseudo-predicates.

3.5.1 *Pseudo-predicates*

Pseudo-predicates are predicates which obligatorily obtain object clitics which, however, do not agree with the object but what appears to be the subject. Some of these predicates behave like verbs, for example, in being circumfixed with *ma…x* for negation (e.g. *għand-* 'have' (57)), while others are not (e.g. *il-* 'be since' (58b)). Just like verbs, some pseudo-predicates take the *-ni* clitic for first person singular while others take *-i*, which is attached to prepositions and nouns (see 3.2 above). The classic example is *għand* 'have' (57), which is clearly historically derived from 'at'.[17] The following illustrates *għand*, which is negated by *ma*…x but takes the *-i* clitic.[18]

(57a)	*Jien*	*għand-i/*	*ma*	*għand-i-x*	*ktieb.*
	I	have-1SG/	NEG	have-1SG-NEG	book.SGM
	'I have/do not have a book.'				

(57b)	*Susan*	*għand-ha/*	*ma*	*għand-hie-x*	*ktieb.*
	Susan	have-3SGF/	NEG	have-3SGF-NEG	book.SGM
	'Susan has/doesn't have a book.'				

Other examples are *għad-* 'be still', *il-* 'be since' and *qis* 'be like'.

(58a)	*Jien*	*għad-ni*	*d-dar.*
	I	still.be-1SG	DEF-house
	'I am still at home.'		

17 See Comrie (2019), chapter 10, section 10.4.

18 For detailed discussions and analyses of pseudo-predicates, see Fabri (1993) and Peterson (2009).

(58b) *Int* *il-ek* *id-dar.*
you be.still-2SG DEF-house
'You have been home for a long time.'

(58c) *Dawk* *qis-hom* *dundjan-i.*
that.PL be.like-3PL turkey-PL
'Those are like turkeys.'

3.5.2 *Non-formal agreement*

Non-formal agreement occurs when the semantics takes over the terms of agreement, overwriting the morphology (see Fabri (1993), (2009) and Borg & Amaira (2020) for a detailed discussion). There are a number of different types of non-formal agreement. Here I briefly discuss two examples.

Example (59b), as opposed to (59a), involves a shift in the conceptualisation from a unit perspective to an aggregate perspective forcing plural marking on the verb and, therefore, (dis)agreement in number.

(59a) *Sandra* *wasl-et.*
Sandra.3SGF arrive-3SGF
'Sandra has arrived.'

(59b) *Sandra* *wasl-u.*
Sandra.3SGF arrive-3PL
'Sandra (and her family/friends) have arrived.'

Note that the subject *Sandra* is third person feminine singular. This agrees with the verb *waslet* in (59a) but disagrees with the verb *waslu* in (59b), which is third person plural. The plural verb forces a reinterpretation of the subject as referring not to one individual but to a group. This phenomenon is generally restricted to spoken discourse.

Examples (60b) and (60c) involve a metonymic interpretation of the noun *bużżieqa* 'balloon'.

(60a) *Dik* *il-bużżieq-a* *nfaħ-t-ha.*
that.SGF DEF-balloon-SGF blow-1SG.PFV-3SGF.DO
'I blew that balloon.'

(60b) *Dak* *il-bużżieq-a* *beda* *j-ibk-i.*
that.SGM DEF-balloon-SGF start.3SGM.PFV 3M.IPFV-cry-SG
'That touchy male started crying.'

(60c) *Dik* *il-bużżieq-a* *bd-iet* *t-ibk-i.*
that.SGF DEF-balloon-SGF start.3SGF.PFV 3F.IPFV-cry-SG
'That touchy female started crying.'

In (60b), the head noun *bużżieqa* 'balloon' in the subject noun phrase is formally feminine singular, as can be seen from (60a); however, unlike (60a), the demonstrative (and the two verbs) in (60b) are masculine singular. This forces a reinterpretation of the noun *bużżieqa*, which is made to refer to a male entity, and thus take on the meaning 'touchy person' instead of 'balloon'. Note that, if the demonstrative and verb are feminine singular, there can still be a shift in the meaning of *bużżieqa,* but this shift comes from the meaning of the verb *tibki*, which requires an animate/human subject. This shows that in (60a) verb meaning and agreement 'conspire' to force a reinterpretation of *bużżieqa*.

4. Conclusion

Grammatical agreement is a phenomenon that involves several core areas of the grammar, in particular morphology, syntax, semantics and pragmatics (discourse). This chapter illustrates the various instances of agreement within various structural domains, and involving gender, number and person distinctions in Maltese. We also explore the various ramifications of this phenomenon within the grammar and discuss cases of 'irregular' agreement, which help to better understand the core phenomena and shed light on the nature of the agreement phenomenon. The next step

is to develop a theory or model of agreement in Maltese, which specifies the role and function of agreement within the grammar, and which can then be incorporated into a general theory of agreement in natural language.

Abbreviations

1, 2, 3	first, second, third person	M	masculine
COLL	collective	NEG	negative
CMP	comparative	NOM	nominative
CS	object case marker	PART	participle
DEF	definite article	PL	plural
DO	direct object	POSS	possessive
F	feminine	PFV	perfective
INDEF	indefinite article	PRPO	prepositional object
IO	indirect object	REFL	reflexive
IPFV	imperfective	SG	singular

References

Bach, Emon (1983) On the relationship between word-grammar and phrase-grammar. *Natural Language and Linguistic Theory* 1, 65–89.

Borg, Albert & Amaira, Mark (2020) Will it agree or not? Syntactic overgeneralization leading to agreement mismatches in Maltese. In: Čéplö, Slavomír & Drobný, Jaroslav (eds.), *Maltese linguistics on the Danube.* Berlin & New York: De Gruyter Mouton, 121–140.

Barlow, Michael & Ferguson, Charles A. (eds.) (1988) *Agreement in natural language: Approaches, theories, descriptions*. Stanford: CSLI.

Borg, Albert & Azzopardi-Alexander, Marie (1997) *Maltese. Lingua descriptive grammar.* London & New York: Routledge.

Camilleri, Maris (2009) *Clitics in Maltese*. Masters Dissertation, University of Malta.

Camilleri, Maris (2011) On pronominal verbal enclitics in Maltese. In: Caruana, Sandro, Fabri, Ray & Stolz, Thomas (eds.), *Variation and change: The dynamics of Maltese in space, time, and society*. Berlin: Akademie Verlag, 131–156.

Camilleri, Maris (2014) The Maltese restrictive relative clause. In: Borg, Albert, Caruana, Sandro & Vella, Alexandra (eds.), *Perspectives on Maltese linguistics*. Berlin: Akademie Verlag, 161–200.

Camilleri, Maris & Sadler, Louisa (2011) Restrictive relative clauses in Maltese. In: Butt, Miriam & King, Tracy Holloway (eds.), *Proceedings of the LFG11 Conference*. Stanford, CA: CSLI Publications, 110–130.

Camilleri, Maris & Sadler, Louisa (2016) Relativisation in Maltese. *Transactions of the Philological Society* 114 (1), 117–145.

Čeplö, Slavomír (2014) An overview of object reduplication in Maltese. In: Borg, Albert, Caruana, Sandro & Vella, Alexandra (eds.), *Perspectives on Maltese linguistics*. Berlin: Akademie Verlag, 201–223.

Čeplö, Slavomír (2018) *Constituent order in Maltese: A quantitative analysis*. PhD dissertation, Prague University.

Chomsky, Noam (1993) *Lectures on government and binding.* 7th edition. Berlin: Mouton de Gruyter.

Chomsky, Noam (2000) Minimalist inquiries. In: Roger, Martin, Michaels, David, & Uriagereka, Juan, (eds.), *Step by step: Essays on minimalism in honor of Howard Lasnik*. Cambridge, Mass: MIT Press, 83–155.

Comrie, Bernard (1989) *Language universals and linguistic typology: Syntax and morphology*. Oxford: Blackwell.

Corbett, Greville (2008) *Agreement*. Cambridge: CUP.

Fabri, Ray (1993) *Kongruenz und die Grammatik des Maltesischen* (Agreement and the Grammar of Maltese). Tübingen: Niemeyer.

Fabri, Ray (1994) The syntax of numerals in Maltese. In: Brincat, Joseph M. (ed.), *Languages of the Mediterranean. Proceedings of the conference held in Malta, 26-29 September 1991*. Malta: University of Malta, 228–239.

Fabri, Ray (2001) Definiteness marking and the structure of the NP in Maltese. *Verbum*, 23(2), 153–172.

Fabri, Ray (2009) To agree or not to agree: suspension of formal agreement in Maltese. In: Fabri, Ray (ed.), *Maltese linguistics: A snapshot. In memory of Joseph A. Cremona (1922–2003). Il-Lingwa Tagħna – Our Language- vol. 1*. Bochum: Brockmeyer, 35 - 61.

Fabri, Ray & Borg, Albert (2002) Topic, focus and word order in Maltese. In: Abderrahim, Youssi, Benjelloun, Fouzia, Dahbi, Mohamed & Iraqui-Sinaceur, Zakia (eds.), *Aspects of the dialects of Arabic today. Proceedings of the 4th Conference of the International Arabic Dialectology Association (AIDA). Marrakesh, Apr. 1-4.2000. In honour of Professor David Cohen*. Rabat: Amapatril, 354–363.

Fabri, Ray & Borg, Albert (2017) Modifiers and complements within the Maltese verb sequence. In: Saade, Benjamin & Tosco, Mauro (eds.), *Advances in Maltese linguistics.* Berlin & New York: De Gruyter Mouton, 67–86.

Lapointe, Steven G. (1985) *A theory of grammatical agreement.* New York: Garland Publishing Company.

Lehmann (1982) Universal and typological aspects of agreement. In: Seiler, Hansjakob & Stachowiak, Franz, Josef (eds.), *Apprehension. Das sprachliche Erfassen von Gegenstaenden. Teil II: Die Techniken unf ihr Zusammenhang in den Einzelsprachen.* Tuebingen: G. Narr, 201–267.

Maas, Utz (2009) Complex predicates in Maltese: From a neo-Arabic perspective. In: Comrie, Bernard, Fabri, Ray, Hume, Elizabeth, Mifsud, Manwel, Stolz, Thomas & Vanhove, Martine (eds.), *Introducing Maltese linguistics*. Amsterdam: John Benjamins,113–132.

Matasović, Ranko (2018) *An aerial typology of agreement systems.* Cambridge: CUP.
Mel'cuq, Igor (2006) *Aspects of the theory of morphology*. Berlin: Walter de Gruyter.
Moravcsik, Edith A. (1978) Agreement. In: Greenberg, Joseph H. (ed.), *Universals of human language- vol. 4. Syntax.* Stanford: Stanford University Press, 331–374.
Peterson, John (2009) Pseudo-verbs. In: Comrie, Bernard, Fabri, Ray, Hume, Elizabeth, Mifsud, Manwel, Stolz, Thomas & Vanhove, Martine (eds.), *Introducing Maltese linguistics*. Amsterdam: John Benjamins, 181–205.
Robins, Robert H. (1971) *General linguistics: An introductory survey*. 2nd edition. London: Longman.
Sande, Hannah (2019) Phonologically determined nominal concord as post-syntactic: Evidence from Guébie. *Journal of Linguistics- vol. 55*, 4, 831–878.
Spagnol, Michael (2011) *A tale of two morphologies*. *Verb structure and argument alternations in Maltese*. Konstanz: Universität Konstanz (Doctoral Dissertation).
Stolz, Thomas (2009) Splitting the verb chain in modern literary Maltese. In: Comrie, Bernard, Fabri, Ray, Hume, Elizabeth, Mifsud, Manwel, Stolz, Thomas & Vanhove, Martine (eds.), *Introducing Maltese linguistics*. Amsterdam: John Benjamins, 133–180.

CONSTITUENT ORDER IN MALTESE: A QUANTITATIVE REEVALUATION[1]

Slavomír Čéplö

Abstract

This paper examines the question of constituent order in Maltese in light of major approaches to it and previous descriptions of Maltese. Using a syntactically annotated corpus (treebank), a quantitative analysis of constituent order in various clause types is performed. This analysis confirms that the default order in Maltese is SVO (with VS in existential clauses as the only exception). Furthermore, it is found that the constituent order in Maltese is quite rigid, more akin to English than – as has been previously argued – to languages with pragmatically determined order.

Dan l-istudju jeżamina l-ordni tal-kostitwenti fil-Malti fid-dawl tal-approċċi ewlenin u tad-deskrizzjonijiet tal-Malti s'issa. Permezz ta' korpus annotat sintattikament (treebank), issir analiżi kwantitattiva tal-ordni tal-kostitwenti f'diversi tipi ta' sentenzi. Din l-analiżi tikkonferma li l-ordni tipika fil-Malti hija SVO (bl-unika eċċezzjoni ta' VS f'sentenzi eżistenzjali). Barra minn hekk, turi li l-ordni tal-kostitwenti fil-Malti hija pjuttost riġida, u aktar tixbah lill-Ingliż milli – kif hemm min argumenta – lil-lingwi, li fihom l-ordni tal-kostitwenti hija determinata b'mod pragmatiku.

1 This paper is a revised and condensed version of chapters 1, 2, 3 and 7 of my dissertation (Čéplö 2018).

1. Introduction

1.1 General

Constituent order, i.e. the order of the verb (V) and its main arguments – the subject (S) and the direct object (O) – within a clause or sentence,[2] is one of the fundamental elements of syntactic description. Its importance is evidenced by the fact that it is often the only piece of information available on the syntax of a language; indeed as Dixon (2009: 73) notes, since most of the world's languages are under-described, it is often the only piece of information on the grammar of a language available. Comprehensive overviews of the world's languages such as *Ethnologue* (Lewis et al. 2016) are the best witness to this. To pick two random examples: the *Ethnologue* entry for Swedish (ISO 639-3 code "swe"), a relatively small but well-described language, lists the following under "Typology":

> SVO; prepositions; noun head final; gender (common, neuter); definite and indefinite articles; passives (active, middle, passive); comparatives; 19 consonant and 17 vowel phonemes; tonal (2 tones).

For Övdalian (ISO 639-3 code "ovd"), also spoken in Sweden, a close relative of Swedish and thus hardly an exotic language, the same section contains only the following:

> SVO; 24 consonants, 9 vowels, 6 diphthongs and 1 triphthong.

The noticeably frequent appearance of constituent order in even the most rudimentary language descriptions is likely due

2 In what follows, I will use the term "constituent order" as defined above. The term "word order" is often used in this sense as well, but for clarity's sake, I will define "word order" as the order of elements within a phrase (e.g. the order of nouns and adjectives) and use it in this sense throughout.

to two factors: first, constituent order is typologically associated with a number of other syntactic and even morphological features and can thus serve as a microcosm of a language's grammar. Secondly, constituent order is one of those properties of a language that are conspicuous (especially when different from what one is used to) and thus seem relatively easily discernible, much like its phonological inventory (again, see the Övdalian example above).

The former is arguably correct, at least to some extent (on which see 1.2 below); the latter, however, is not entirely so and the answer to the question of what the constituent order of a particular language is will almost invariably be a complex one. In this paper, I will attempt to provide it for Maltese, considering its context within both general and Maltese linguistics.

1.2 Constituent order and typology

The undoubtedly most influential work on constituent order in modern linguistics is Joseph H. Greenberg's 1963 paper titled *Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements* (cited from the second edition, Greenberg 1966). Greenberg expanded relatively trivial observations on how languages differ in the order of "modifying or limiting elements" (Greenberg 1966: 76) into a full-fledged typological classification of languages based on a list of so-called universals. The fundament on which these rest is his basic order typology: Greenberg takes the observation that "languages have several variant orders but a single dominant one" (Greenberg 1966: 76) to its logical conclusion and establishes a six-way typology of dominant orders of subject, verb and object: SVO, SOV, VSO, VOS, OSV and OVS. He immediately notes, however, that three of those – VOS, OSV and OVS – "do not occur at all, or at least are rare" (Greenberg 1966: 76) and proceeds to draw from this his first universal:

> Universal 1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.

Greenberg combines the remaining three configurations – referred to as Type I (VSO), Type II (SVO) and Type III (SOV) – with two additional binary criteria (whether a language has prepositions or postpositions and whether an adjective of quality follows the noun it modifies or precedes it) and investigates the correlations between these syntactic properties in a sample of 30 languages (Greenberg 1966: 74-75):

Basque, Serbian, Welsh, Norwegian, Modern Greek, Italian, Finnish (European); Yoruba, Nubian, Swahili, Fulani, Masai, Songhai, Berber (African); Turkish, Hebrew, Burushaski, Hindi, Kannada, Japanese, Thai, Burmese, Malay (Asian); Maori, Loritja (Oceanian); Maya Zapotec, Quechua, Chibcha, Guarani (American Indian).

Using these correlations as the starting point, Greenberg postulates 45 implicational universals, 15 of which relate to constituent order or at least the position of the verb and its arguments, including question words.

Greenberg's universals were met with almost immediate acceptance and despite substantial criticism (on which see below) and some empirical evidence to the contrary (like the case of OVS order in Hixkaranya described by Derbyshire 1977), Greenberg's six-way typology continues to be the dominant paradigm in the cross-linguistic study of constituent order variation. Works like Payne (1997: 71-74), Song (2011b), the *Ethnologue* (see the entries above) and *The World Atlas of Language Structures* (WALS; Dryer and Haspelmath 2013) are but a few of the most prominent examples of Greenberg's enduring legacy.

1.3 The problem of 'basic' constituent order

As with any new paradigm, criticisms of Greenberg began to appear almost immediately. One of the primary issues that emerged as a major point of contention is the problem of basic (default) word order. Greenberg's original formulation of his universal does not actually define what qualifies as 'basic', merely assumes it: "If a language has verb-subject-object as its basic word order in main declarative clauses..." (Greenberg 1966: 74). Greenberg is aware that this presupposes, at the very least, the existence of a subject-predicate structure in all languages under investigation. He acknowledges the problems with this assumption, but proceeds without resolving this issue, since doing so would have "prevented me from going forward to those specific hypotheses, based on such investigation, which have empirical import and are of primary interest to the non-linguist" (Greenberg 1966: 74). In other words, Greenberg was primarily interested in the universals (and their correlations) and was willing to sacrifice accuracy in determining the basicness of a particular constituent order in a particular language to achieve his goal.

This is obviously a problem and one that is related to a larger issue in linguistics: if the ostensible goal of linguistics (or at least its descriptive and typological branches) is to provide a description of one or more languages, then the primary question becomes what it is one is actually describing. For example, a linguist who is a native speaker of a language could base their description of that language on their own knowledge. Such an approach to linguistic description, commonly referred to as introspective or intuitive (Itkonen 2005), is (or at least was) typical for generative linguistics; one infamous instance involves Noam Chomsky arguing that the English verb 'perform' cannot take mass nouns as objects[3] and insisting he is correct because "I am a native speaker of the English language" (Harris 1995: 97). Whether such an

3 He was, as is often the case with such pronouncements in general, wrong.

approach to the analysis of anything is truly scientific is best left for another time and venue; what matters is that there is an obvious practical issue with this approach: what happens if another native speaker disagrees, as one immediately has in the case described above (Harris 1995: 97)?

The only other option available to a linguist is to collect data, i.e. the empirical approach. Within modern linguistics, there are two major ways of doing this: the first one is elicitation, which essentially involves asking many native speakers, thus hopefully at one point arriving at a consensus or at least clearly defined variation. This is a tried and true method, but it often brings with itself not only practical challenges (e.g. how much is many, how one gets cooperative respondents etc.), but also entails problems of epistemological nature: human beings have all types of ideas and preconceptions about language; chances are, therefore, that asking them about their language and their use thereof will yield information that is not objective, reflecting the respondents preconceptions, rather than the actual linguistic reality.

The other route to take is to use a corpus, i.e. a collection of texts (whether they originated in writing or they came about as transcriptions of speech) in a particular language. The corpus approach, often taken to be synonymous with the empirical approach, is nothing new in principle – grammarians and lexicographers have been using collections of texts to do their work for centuries. Modern corpus linguistics, however, does differ from those in two ways, both thanks to the relatively recent advances in computing: first, modern-day corpora are by orders of magnitude larger than those available to anyone in history. The size of corpora, along with the fact that texts typically contain spontaneously produced language, is the main advantage of corpus linguistics over elicitation, as it eliminates the epistemological issues associated with the latter described above. Second, the use of computers to store and query those corpora has inevitably put

large emphasis on quantitative measures, especially frequency, which has led to some surprising insights, such as the Menzerath's Law (Milička 2014).

This, once again, is nothing new in principle: Greenberg's work is, after all, all about statistics. The problem described above lies in the fact that the proper statistical considerations (sampling, sample size, representativeness etc.) are only applied to the universals, not to the analysis of individual languages. Many of Greenberg's successors and critics have attempted to correct this, but Matthew Dryer is by far the most successful and thus most influential. Dryer's work on constituent order typology began as a criticism of Greenberg's sampling methods and a test of hypotheses raised by Greenbergian universals (Dryer 1989b) and included a large follow-up study of the universals using a larger and more balanced sample of languages (Dryer 1992). This work led Dryer to renounce Greenbergian six-way typology and propose a new typology, based on two independent but interacting binary parameters, SV/VS and VO/OV (Dryer 1997, Dryer 2013b). Dryer lays out a complex case for this, the chief arguments being that "some word order parameters correlate with both the order of the object and the verb and with the order of the subject and the verb" (Dryer 2013b: 295) and that a typology based on these two parameters is more fundamental than the six-way typology, as it is "based on clause types that occur much more frequently" (Dryer 1997: 70). The latter illustrates Dryer's focus on frequency as an important element in linguistic description and explanation: Dryer recognizes that "speakers store grammatical knowledge independent of frequency", but argues that "frequency plays a pervasive role in explaining why languages — and grammars — are the way they are" (Dryer 2013b: 292). Consequently, Dryer's concept of basic order is based solely on frequency where, admirably, Dryer is aware of the inherent dangers of inadequate sampling (Dryer 1997: 72, italics in the original):

> If a particular order is more common in most or all texts, then we can justifiably describe that order as most frequent. If no order is most frequent over most texts, however, or if the order varies from genre to genre or text to text, we should probably not describe any particular order as the basic order (in the sense of most frequent order) and we should say that the language is one that lacks a basic word order [...]. In short, while it may be relatively easy to identify a most frequent order in a single text or in a small body of texts, it is necessary to examine a wide variety of texts before one can decide with confidence that a particular order is most frequent in the language *as a whole*.

In typological studies of word and constituent order, Dryer's work has become the standard reference, as evidenced not only by his contribution to general discussions on the state of the question (see the special issue of *Linguistic Typology 15*), but also his authorship of chapters on word order in such overviews of language typology as Shopen 2007 (Dryer 2007) or WALS (Dryer 2013a and 2013c). And while the latter work also uses Greenbergian six-way typology in its description of constituent order typology (though not exclusively), it is here that Dryer provides the ultimate definition of basic or – in Dryer's terminology – dominant order defined in terms of frequency (Dryer 2013a):

> The expression *dominant order* is used here, rather than the more common expression *basic order*, to emphasize that priority is given here to the criterion of what is more frequent in language use, as reflected in texts. ... The rule of thumb employed is that if text counts reveal one order of a pair of elements to be more than twice as common as the other order, then that order is considered dominant, while if the frequency of the two orders is such that the more frequent order is less than twice as common as the other, the language is treated as lacking a dominant order for that pair of elements. For sets of three elements, one order is considered dominant if text counts reveal it to be more than twice as common as the next most frequent

order; if no order has this property, then the language is treated as lacking a dominant order for that set of elements.

This definition (applicable to both word order and constituent order, and both pairs and triads) is specific, empirically founded, without any theoretical baggage, cross-linguistically applicable, and clearly actionable (step 1: get texts; step 2: count); as such, it constitutes a significant improvement to previous definitions of "basic" constituent (and word) order; it will therefore be adopted in what follows under the name Dryer's 2:1 method.

1.4 The problem of 'free' constituent order

The typological classification of languages by basic constituent order assumes that such a basic order exists in all languages. It has, however, long been known that there exist languages with seemingly endless variation in their constituent order, also known as "free word/constituent order" languages; in fact, it is probably the oldest classification of languages by constituent order, dating at least as far back as Weil (1844: 25). Weil's observations focused on flexibility of the order of constituents in classical Greek and Latin compared to the relative rigidity in modern languages such as French and German and were thus somewhat of a surprising revelation. To other linguists, such as those of the Prague Linguistic Circle almost a century later, the fact that some languages are very flexible when it comes to constituent order was no surprise, since their own native language – Czech – was one. The relatively free constituent order in Czech led Vilém Mathesius to the fundamental insight that in some languages, constituent order and pragmatics (i.e. the context in which a sentence is produced and the purpose for which it is produced) are intrinsically linked and "[t]he functional analysis of a sentence must be juxtaposed to its formal analysis" ("Aktuální členění věty je třeba klásti proti jejímu členění formálnímu." Mathesius

1939: 171; see Firbas 1992: 22 for the English terminological choice). Expanding on previous work by Weil (1844) and von der Gabelentz on the distinction between grammatical subject and "psychological subject" ("das psychologische Subjekt", von der Gabelentz 1869: 378), Mathesius establishes a two-way division of sentence in terms of its communicative effect: the "theme", defined as "a thing about which we assert something" ("to, o čem něco tvrdíme", Mathesius 1961: 91) and "what we say about the theme is the nucleus or the enunciation" ("to, co o základu tvrdíme, je jádro výpovědi neboli vlastní výpověď", Mathesius 1961: 92). This division, for which Mathesius' successors (Firbas 1957) established the terms "theme" and "rheme", is the cornerstone of what has become known as the Functional Sentence Perspective (FSP). And while FSP as a theory of communication is largely unknown outside of Czech linguistics, its foundational works by Mathesius (1961 in its English translation) and Firbas (1964) are credited with establishing the subfield of information structure (Féry and Ishihara 2016b: 3). Its basic terminology, redressed and redefined multiple times – typically as 'topic' and 'comment' or 'topic' and 'focus' – and its fundamental ideas like context-boundness (Krifka and Musan 2012) have become a firm part of modern linguistic terminology (Féry and Ishihara 2016a).

Mathesius was far from the only one to notice the relationship between constituent order and pragmatics. Even Chomsky, despite his focus on structural description formulated as transformation rules, recognizes the importance of pragmatics (or, in his words, "stylistic factors"; Chomsky 1965: 11) for the variation of constituent order, noting that "grammatical transformations do not seem to be an appropriate device for expressing the full range of possibilities for stylistic inversion" (Chomsky 1965: 126). He resolves this conundrum by claiming that the rules of pragmatically determined variation in constituent order "are not so much rules of grammar as rules of performance" and while interesting, they have "no apparent bearing, for the moment, on the theory of grammatical

structure" (Chomsky 1965: 127). The moment in question did not last long and soon generativist works began to appear dealing with "the annoying problem that languages differ from one another" (Carnie 2013: 27) in the ordering of the constituents. John R. Ross' 1967 PhD dissertation devotes some attention to the problem of free word order in Latin and other languages in the context of node deletion or tree pruning, i.e. reducing the complexity of sentences generated by existing theories of generative grammar (Ross 1967: 41). In the analysis of the various possible configurations of constituents and even components of noun phrases in Latin, Ross proposes the Scrambling Rule (Ross 1967: 75) which permits the seemingly unlimited surface variation of words in Latin sentences. Since Ross's day, two approaches have developed to account for scrambling: the base-generation approach argues that variation in constituent order is a syntactic phenomenon, i.e. it is generated randomly at the D-structure level (Corver and van Riemsdijk 1994b: 1). The distinction made here is between configurational languages which do not allow this random generation of constituents and non-configurational languages (also termed "flat languages" by Hale 1983: 10, since they do not have a unitary Verbal Phrase) which do. In contrast, the movement approach (Corver and van Riemsdijk 1994b: 2) explains variation in constituent order by different types of movements, such as object shift (e.g. Broekhuis 2008 for Germanic languages) or VP fronting (Zubizarreta 1998). Both approaches have produced much literature (see Corver and van Riemsdijk 1994a for an overview), but so far, without any consensus in sight.

While the generativist discussion of scrambling seems to be dominated by the base-generation and movement approaches, there is still a third school of thought harkening back to Chomsky 1965 and Ross 1967 which considers constituent order variation from the point of view of pragmatics. This school, best represented by Kiss (1995a), has surveyed a number of languages very different from Standard Average European (Kiss 1995b: 4) and observed that "the structural role that the grammatical subject plays in the

English sentence may be fulfilled by a constituent not restricted with respect to grammatical function or case in other languages" (Kiss 1995b: 3). In simple terms, this school of thought argues that languages fall into two groups: subject-prominent languages where the surface constituent order is Subject – Verbal Phrase and topic-prominent languages, where the place of the Subject can be taken by an arbitrary element bearing a particular discourse (or pragmatic) function (Kiss 1995b: 4). These languages are termed discourse-configurational and their fundamental properties are as follows (Kiss 1995b: 6):

> A. The (discourse-)semantic function 'topic,' serving to foreground a specific individual that something will be predicated about (not necessarily identical with the grammatical subject), is expressed through a particular structural relation (in other words, it is associated with a particular structural position).
> B. The (discourse-)semantic function 'focus,' expressing identification, is realized through a particular structural relation (that is, by movement into a particular structural position).

One crucial aspect of the theory behind discourse configurationality is the empirical distinction between categorical and thetic statements (Kiss 1995b: 7-8). The distinction is based on Marty's (1897) observation that there exist two types of sentences: those that do not express judgments (in the philosophical sense), like interrogative or imperative sentences (Marty 1897: 189), and those that do. Furthermore, the latter group can be divided into two types: the first type is referred to as compound or categorical judgments which actually contain two judgments, one about the existence of the subject and the other about a property of the subject. The second type is referred to as pseudo-categorical or thetic judgments (Marty 1895: 298) and they contain a single judgment only; these typically include existential, impersonal and universal sentences (Kiss 1995b: 7). A language can be discourse-

configurational with the property A only if it differentiates between categorical and thetic sentences syntactically.

Kiss goes on to argue that while sometimes properties A and B go hand in hand, they are not interdependent and so some discourse-configurational languages can display only type A characteristics, whereas others only show the type B properties (Kiss 1995b: 6). It should be noted, however, that while the fundamentals of this subset of generativist theory are framed in terms of pragmatic function, much of the explanation offered by its proponents still depends on movements (Choe 1995), such as the Focus Movement (focalization) and the Topic Movement (topicalization). And as with literature on scrambling, there seems to be no consensus in generativist literature on the general properties and nature of discourse configurationality. The term, however, is often used as nearly synonymous with "pragmatically determined word/constituent order" or its equivalents and, by extension, "free word/constituent order".

2. Studies of constituent order in Maltese

For a numerically small and geographically and culturally marginal language, Maltese boasts a remarkably long and rich tradition of scholarly interest. This is evidenced, inter alia, by the fact that the first grammatical description of Maltese worthy of the name, de Soldanis' 1750 *Nuova scuola di grammatica per agevolmente apprendere la lingua punica – maltese* (published in de Soldanis 1750), predates the first actual printed book in Maltese (Francesco Wzzino's translation of the Catholic Catechism titled *Taġhlim Nisrani* published in Rome) by two years. In the intervening 270 years, many grammars of Maltese have been written, some of which addressed the question of constituent order in one form or another. A detailed analysis would require more space than is available here (and in any case, I have provided it in Čéplö 2018: 31-49), so Table 1 below summarizes their findings.

Work	Classification
Vella 1831: 224-225	- SV - VS in relative clauses
Sutcliffe 1936: 210	- VSO with variation "for euphony or emphasis" - VS in subordinate clauses
Aquilina 1959: 341	- SVO as the default - VS in "emphatic or high-flown literary language" and in subordinate clauses
Vella 1970: II.98	- VS "as is the Semitic custom" - SV "[d]ue to foreign influence"
Krier 1976: 79	- SV with "liberté de position est due à la mise en valeur stylistique (variation is due to stylistic emphasis)"
Kalmár and Agius 1983: 336-337	- SV - pragmatically determined VS
Fabri 1993: 7, 131	- "relative freie Wortstellung (a relatively free word order) " - "eine konfigurationale Sprache (a configurational language)" - considerable variation
Borg and Azzopardi-Alexander 1997: 57	- SVO(I) as "neutral order"
Fabri and Borg 2002: 362	- SV - VS with stress on V - SVO - OVS with stress on O
Fabri 2010: 793-794	- "a topic-oriented language" - "relatively free" - "SVO"
Borg and Fabri 2016: 417	- "a discourse configurational ... language, especially in its spoken form"

Table 1: Overview of previous descriptions of constituent order in Maltese

As Table 1 shows, two constant themes are interwoven throughout the history of the study of Maltese constituent order:

First, there is the question of what is the default (unmarked, basic, dominant) constituent order in Maltese. This has been answered in at least two different ways: verb-first, as argued by Sutcliffe 1936 and Vella 1970; or subject-first, as described by Aquilina 1959, Kalmár and Agius 1983, Borg and Azzopardi-Alexander 1997 and others.

The other theme is that of classifying Maltese constituent order as 'free' (e.g. Fabri 1993: 7, 131 and Fabri 2010: 793), including its near-synonyms like "discourse-configurational" (Fabri and Borg 2002, Borg and Fabri 2016) and "topic-oriented" (Fabri 2010: 793, Fabri and Borg 2017: 83). All those terms describe Maltese as a language where "constituent order, at sentence level is strongly influenced by pragmatic factors, in particular topic and focus, contrast and emphasis, more than by syntactic factors" (Fabri and Borg 2017: 83). In this context, a number of authors note a great deal of variation in Maltese constituent order (Sutcliffe 1936: 211, Krier 1976: 79, Fabri and Borg 2002) and attempt to account for it (Borg and Azzopardi-Alexander 1997, Fabri and Borg 2002).

Additionally, a number of works (e.g. Borg and Azzopardi-Alexander 1997, Borg and Azzopardi-Alexander 2009 and Čéplö 2014) devote a significant amount of attention to topicalization of direct and indirect objects, i.e. the placement of the object before the verb, typically also accompanied by a resumptive clitic and a phonological break. This phenomenon, which according to Borg and Azzopardi-Alexander (1997: 126) "is such a wide spread characteristic of Maltese, that it even features in Maltese English", is related to both the question of the default constituent order in Maltese, as it at the very least assumes VO as the default, as well as to the question of the influence of pragmatic factors on the same.

All these analyses can be shown to have serious shortcomings: for the question of the default (unmarked, basic, dominant), the chief one is obviously the lack of general agreement. Additionally, there are multiple methodological issues, ranging from the lack of a meaningful definition of "default (unmarked, basic, dominant)" constituent order, through the lack of detailed studies on clause-type level (with Borg and Azzopardi-Alexander 1997 as sole attempt to do so in a systematic manner), all the way to the fact that most such studies have been introspective at best, impressionistic at worst. Even those that employed some sort of empirical approach (which is the case for Krier 1976 and Kalmár

and Agius 1983) did so more than imperfectly, rendering their conclusions tentative at best. Much of this also applies to works which describe Maltese constituent order as free or pragmatically determined; additionally, these have problems of their own. And so for example even those studies that provide a detailed account of the possible variation based on pragmatic (information structure) factors (Borg and Azzopardi-Alexander 1997, 2009; Fabri and Borg 2002) essentially only described potentiality, i.e. what options are available to speakers of Maltese, but did not (except in the broadest terms, e.g. Borg and Azzopardi-Alexander 1997: 126) provide a description of how those possibilities are instantiated.

In what follows, I will try to remedy those shortcomings by using an empirical analysis.

3. Quantitative analysis of constituent order in Maltese

3.1 Methodology and data

Having reviewed the major ways of analyzing and classifying constituent order (section 1) and how they have been applied to the analysis of constituent order in Maltese (section 2), we can now proceed with the actual analysis. For the methodology, I will use Dryer's 2:1 method to analyze both the Greenbergian six-way classification, as well as Dryer's two-way classification.

As advertised above, the analysis I am about to conduct is empirically founded, i.e. corpus-based. Such an analysis, however, requires a syntactically annotated corpus (also known as a treebank); neither of the two large corpora already available for Maltese (MLRS and *bulbulistan*; Gatt and Čéplö 2013) contain such annotation. The solution is to compile a treebank of Maltese, which I have done. For the annotation scheme, I have chosen that employed by the Universal Dependencies project (UD; Zeman,

Nivre, Abrams et al. 2020), a de facto standard in syntactic annotation of corpora for NLP purposes. As with most languages, the UD annotation scheme had to be adapted to Maltese. The process is somewhat complicated as it amounts to compiling a sketch of Maltese syntax and there is, sadly, no space here for the full description; those interested are welcome to consult my dissertation (Čéplö 2018: 83-171).

What does require further elaboration, however, is the extent and composition of the Maltese UD treebank (henceforth: MUDT; at the time of writing in the version 2.7, hence MUDT v2.7). As this is the first effort in compiling a syntactically annotated corpus of Maltese, the vast majority of the annotation would have to be done manually, and so a balance had to be found between the desire to end up with as much data as possible and the practicality of what could be achieved with a manageable amount of effort within a reasonable time frame. In the final count of 44,162 tokens in 2074 sentences, MUDT is comparable to UD treebanks for such languages as Vietnamese, Wolof or Hungarian, each of which has many more speakers than Maltese.

The issue of the composition of MUDT is directly related to its size and by extension to the problem of whether corpus data accurately reflect the language under investigation. In corpus linguistics, this is a critical issue and several solutions have been adopted (McEnery and Hardie 2011: 6-10). Considering the fact that the treebank had to be drawn from the existing corpora which are opportunistic by nature (i.e. based on the "we take all we can get" principle) and are composed of roughly four different genres, the solution I adopted for MUDT was to create a balanced treebank where the four genres – or text types – would be represented more or less equally in terms of sentence counts. This ensures that any description of Maltese based on MUDT is not just a description of a single genre, say, the journalistic language, which is very well known to differ substantially from other genres (Suter 1993).

Table 2 summarizes the composition of MUDT. The text type and subtype descriptions are self-explanatory, save perhaps for

the 'quasi-spoken' text type: I have chosen this label because while the texts in that group do originate from spoken language (interviews and parliamentary debates), they have undergone some form of editorial processing and as such cannot be considered transcriptions of speech.

Text type	**Subtype**	**Sentence count**
newspaper	news	239
	op-eds	240
	Subtotal	*479*
quasi-spoken	newspaper interviews	280
	parliament: debates and Q&A	294
	Subtotal	*574*
fiction	short stories	246
	novel chapters	251
	Subtotal	*497*
non-fiction	humanities	249
	science, encyclopedic and instructional	275
	Subtotal	*524*
	Total	**2074**

Table 2: The composition of MUDT v2.7 by genre

3.2 The analysis

Having established our methodology and the data set, we can now proceed to data collection and analysis. For the former, I have opted to import MUDT (in its most recent version v2.7) into an instance of the corpus management software ANNIS3 (Krause and Zeldes 2016) available at https://bulbul.sk/annis-gui-3.6.0 (item *MUDT_v27*). I then ran a number of queries to obtain the data in question, such as this one:

tok ->dep[deprel=/nsubj/] tok & #1 ->dep[deprel=/obj|obl:arg/] tok & #3 .* #2 .* #1

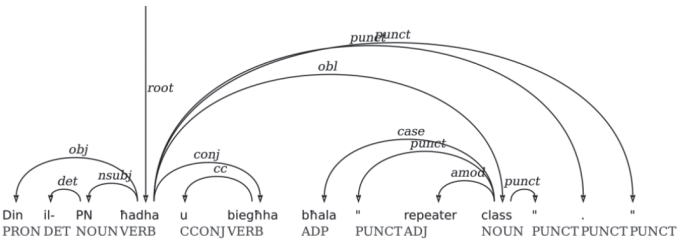
The query searches for:

- any token (first *tok*)[4]
- which has a nominal subject token (*dep[deprel=/nsubj/] tok*) as a dependent
- while at the same time (first *&*), it (*#1*) also has an object or a non-canonical object[5] token (*dep[deprel=/obj|obl:arg/] tok*) as a dependent,
- with the added condition (second *&*) that the three tokens must appear in a specified order, i.e. the *obj|obl:arg* (*#3*) token, *nsubj* token (*#2*), the first token (*#1*) and the.

In other words, this query will retrieve all OSV clauses, such as the one in (1).[6]

(1)	*Din*	*il-*	*PN*	*ħadha*		*u*
	this.F	DEF	PN	take.PAST.3SGM-ACC.SGF		and
	biegħha			*bħala*	*"repeater class".*	
	SELL.PAST.3SGM-ACC.SGF			as	"repeater class"	

'This the PN took and sold it as "repeater class".'



[MUDT v2.7, file 22_02J03]

4 Since only verbs or pseudoverbs can have both a subject and an object, the first token will always be one of these parts of speech; we could specify the parts of speech we're looking for directly, e.g. by replacing the first *tok* with *pos = /VERB/.*

5 See Čéplö 2018: 127-128. These include, for example, prepositional objects, such as the objects of the verb *nduna* "to notice" introduced by the preposition *b'*.

6 In the following examples, I will include punctuation with the glossed word, whereas in dependency graphs, punctuation is considered a separate token.

Using this and equivalent queries, the following data was obtained:[7]

Configuration	**Count**	**%**
SVO	445	94.08%
SOV	0	0.00%
VSO	3	0.63%
VOS	11	2.33%
OSV	3	0.63%
OVS	11	2.33%
Total	*473*	*100%*

Table 3: Constituent order in MUDT v2.7 – Greenbergian analysis

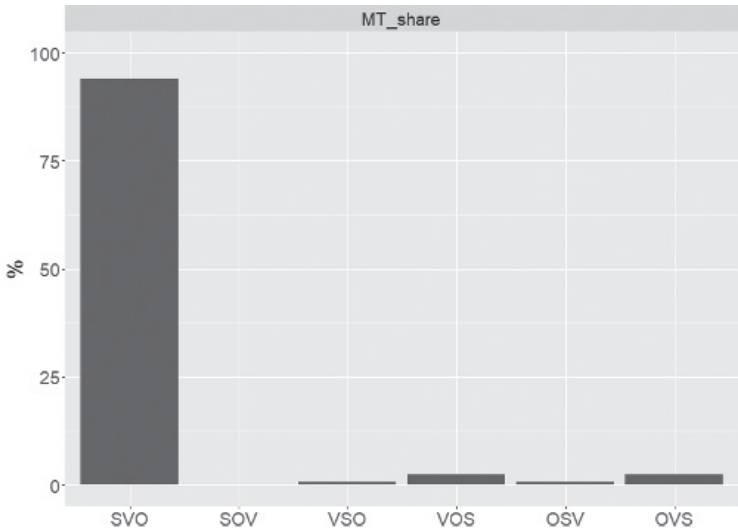


Figure 1: Constituent order in MUDT v2.7 – Greenbergian analysis

7 The data and code used to produce the analysis below can be downloaded from https://bulbul.sk/jms2020.

Configuration	**Count**	**%**
VO	1830	95.11%
OV	94	4.89%
Total V+O	*1924*	*100%*
SV	1697	76.34%
VS	526	23.66%
Total S+V	*2223*	*100%*

Table 4: Constituent order in MUDT v2.7 – Dryerian analysis

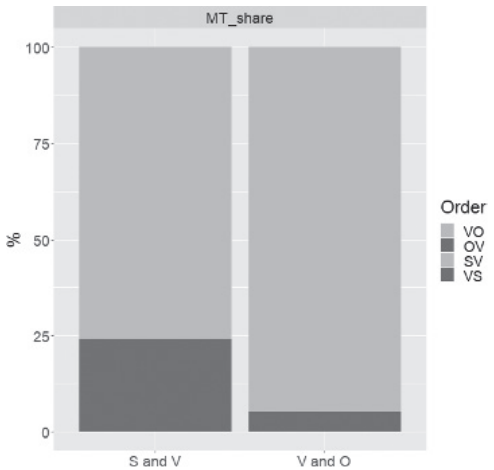
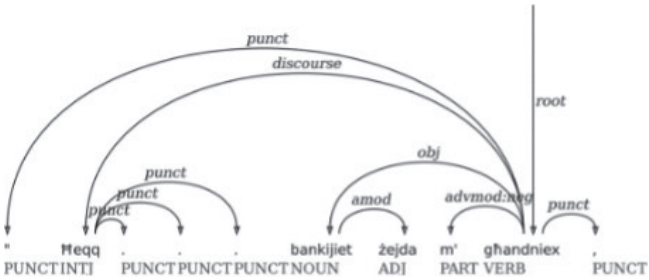


Figure 2: Constituent order in MUDT v2.7 – Dryerian analysis

This data provides a clear picture of both the Greenbergian and the Dryerian classifications of Maltese as, respectively, an SVO and SV/VO language. In fact, Dryer's 2:1 need not even be employed; in all cases, the dominant configuration occurs at least three times as often as the other one.

Of particular interest here is the share of OV clauses, like the one in (2).

(2) “Ħeqq … bankijiet żejda m’ għandniex,
INTJ … bench-PL additional-PL NEG have.PRES-1PL-NEG
‘Yeah … we don’t have any more benches,’



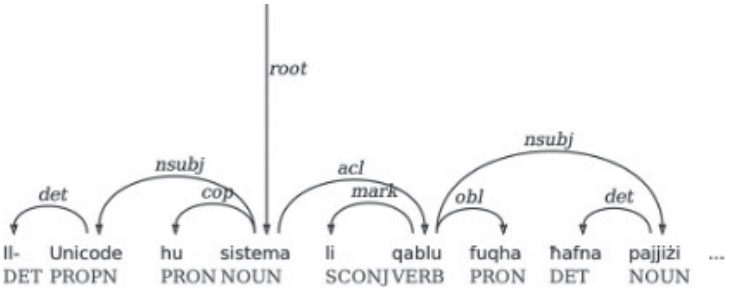
[MUDT v2.7, file 49_03F09]

As noted in section 2 above, these constructions have been described as “a wide spread (sic) characteristic of Maltese” (Borg and Azzopardi-Alexander 1997: 126). And yet in MUDT v2.7, only <5% of all direct objects fall into that group (and that is assuming that all of them represent topicalization which is – what Borg and Azzopardi-Alexander refer to – which is far from certain), a figure which certainly does not represent a widespread phenomenon.

Furthermore, the data offers a clear case for Dryer’s SV/VS and VO/OV typology over the Greenbergian one: with MUDT v2.7, Greenbergian typology only has 487 data points to work with; using Dryerian typology, the data sample expands four-fold for both subjects (2223 total) and objects (1924).

This analysis is of course a rough one and can be refined. One way to do it would be to consider the full spectrum of clause types. Those can be first divided into main and subordinate clauses, which come in several types; of special interest here would be relative clauses (or *acl* in the UD nomenclature, see example 3 below) and adverbial clauses (*advcl*), which some authors (see Table 1 above) described as having VS as the default order.

(3) *Il-* *Unicode* *hu* *sistema* *li* *qablu*
DEF Unicode HE system COMP agree.PERF-3PL
fuqha *ħafna* *pajjiżi* ...
on-3SGF many country.PL …
'Unicode is a system that many countries agreed on…'



[MUDT v2.7, file 57_04N11]

Figure 3 below plots the data for both clause types in MUDT v2.7.

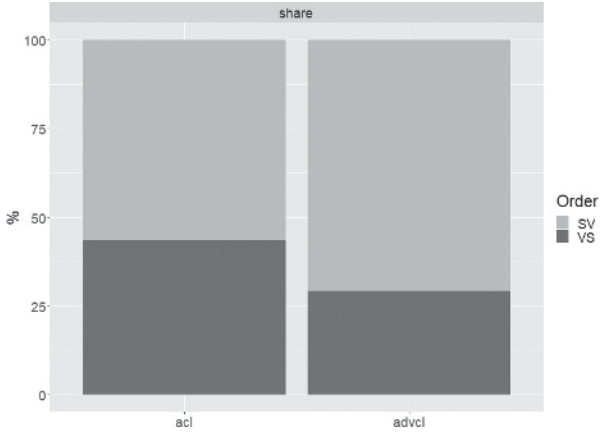


Figure 3: Constituent order in MUDT v2.7 – acl and advcl

As is evident from the plot, the dominant order for *advcl* is SV; no dominant order can be established for *acl* with the distribution of both configurations nearly equal. My preliminary investigation suggests that the VS order in *acl* is positively associated with the heaviness of the subject (i.e. its syntactic complexity and length),

while the SV order is positively associated with the clause length (Čéplö 2018: 199-203). However, more data and a more detailed analysis are required to provide a definitive answer.

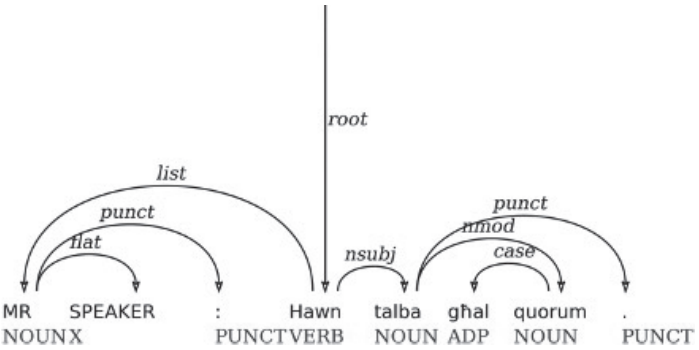
Another subdivision of clause types is that by the word class of the predicate (or root in the terminology of the dependency grammar). Greenbergian analysis is obviously limited to transitive verbs (and pseudo-verbs) only; Dryerian analysis can also take into account intransitive verbs, as well as copular clauses and other clause types. A preliminary analysis (Čéplö 2018: 218-225) has revealed that there is one clause type (defined by the word class of its root) where the default order is VS: existential clauses. These are clauses with the pseudo-verb *hemm* "there is" (and its synonym *hawn*) as the root/predicate, as in (4).

(4)	*MR*	*SPEAKER:*	*Hawn*	*talba*	*għal*	*quorum.*
	Mr	Speaker	EXIST	request	for	quorum

'Mr Speaker: There is a request for quorum.'

Figure 4 plots the distribution of the two possible configurations in existential clauses in MUDT UD v2.5.

This finding confirms an observation by Kalmár and Agius regarding the Maltese constituent order (Kalmár and Agius 1983: 343-344), and also a general cross-linguistic trend: as has been noted on many occasions (e.g. Givón 2001: 257), VS appears to



[MUDT v2.7, file 38_02P06]

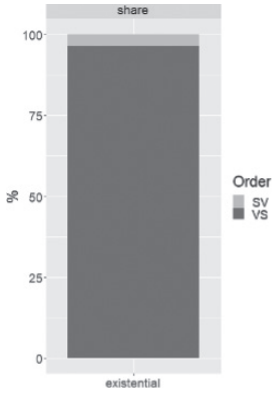


Figure 4: Constituent order in MUDT v2.7 – existential clauses

be the preferred order in existential clauses even in languages which otherwise show clear preference for SV. There are various explanations for this, for which there is little space here. For the purposes of this chapter, it suffices to conclude that Maltese is one of those languages where VS order is the dominant one in existential clauses and at the same time, existential clauses are the only clause type (defined by root) which exhibits this particular configuration as the dominant one.

Even such fine-grained analysis is far from the complete picture of Maltese constituent order, let alone its relationship to clause structure, complex sentence structure, word order, verbal valency and many other problems in Maltese syntax. It is but the first step, and the data provided by the treebank can be used to expand on it and to accomplish much more.

3.3 The problem of 'free' constituent order revisited, or: a két fadatbázis regénye

As we have seen in section 2, Maltese has repeatedly been described as a discourse-configurational language, either explicitly (Fabri and Borg 2002 and Borg and Fabri 2016, both citing Kiss 1995a), or implicitly: Fabri 2010 and Fabri and Borg 2017 describe Maltese

as "a topic-oriented language" (Fabri 2010: 793, see also the almost identical phrasing in Fabri and Borg 2017: 83). Considering the imprecise nature of the terminology, I take this to be a synonym of "topic-prominent language" (Kiss 1995b: 4-5), a term which in the strictest sense designates a subset of languages falling under the "discourse-configurational" umbrella, the so-called type A discourse-configurational languages, where any topicalized constituent can assume the preverbal position typically reserved for the subject (Kiss 1995b: 6-7). In type B discourse-configurational languages, focus-prominent languages, the same is true of focus (Kiss 1995b: 15-24); discourse-configurational languages can be type A, type B or both, depending on the interaction between topic and focus and on inter-language variation. Those works that describe Maltese as discourse-configurational do not elaborate on that particular aspect of this property, but judging from description of focus provided by Fabri 1993 and Fabri and Borg 2002, if Maltese is a discourse-configurational language, it is both type A and type B. This, however, is ultimately irrelevant: Maltese has been described at least twice as discourse-configurational without any elaboration or qualification and it is this description that is the focus of this section.

The framework-dependent reasoning behind this classification is not of interest here. What is, however, is the classification itself, i.e. the claim that Maltese is a discourse-configurational language; more specifically, what I want to focus on is the fact that this claim can be (to some extent) tested. The line of thinking that leads me here is the following:

1. Hungarian is considered the paragon of a discourse-configurational language (cf. Kiss 1995a), i.e. a member of a class of languages defined by a shared property involving constituent order.
2. Maltese has also been described as a discourse-configurational language.

3. Ergo, if one were to investigate the distribution of constituent order configurations in both, one would find that it is at the very least quite similar.

One might also expect that in any discourse-configurational language (and thus both Maltese and Hungarian under assumptions 1 and 2 above), the distribution of SV and VS on one hand and VO and OV on the other would be approximately the same, i.e. 50-50 for both pairs. This is, of course, not realistic, as the theory behind the classification of discourse-configurational languages makes clear: the ordering of constituents is not random[8] but based on pragmatic (and possibly other) criteria. Additionally, the subject is more likely to be the topic (as there is a "close correspondence between the topic and the grammatical subject", Kiss 1995b: 10) and in any case, there are inter-language differences in how far discourse-configurationality goes. Nevertheless, the hypothesis above stands and with the Maltese UD v2.7 (Zeman, Nivre, Abrams et al. 2020) and Hungarian UD v2.5 treebanks (Zeman, Nivre, Abrams et al. 2019), there is a way to test it quantitatively.[9]

To conduct the actual analysis, I replicated the queries used in section 3.2 for both the Maltese UD v2.7 and the Hungarian UD v2.5 treebank. The data obtained is plotted in Figure 5 below.

The data sets underlying these two plots are, needless to say, not the same or even similar. To employ Dryer's 2:1 method (see section 1.3), two different classifications would have to be applied here: Maltese (as represented in UD v2.5) is a language with SVO as the dominant constituent order; Hungarian (as represented in UD v2.5) is a language with no dominant constituent order.

8 On the other hand, both Maltese (Fabri 2010: 793) and Hungarian (Puskás 2000: 41) have been described as having "free word order", so a case could be made that the constituent order in such languages is indeed random (in statistical terms).

9 Hence the subtitle of this section, best translated as "a tale of two treebanks". Having failed to find a commonly used (or indeed any) Hungarian translation of "treebank", I came up with my own, a portmanteau of *fa* "tree" and *adatbázis* "database".

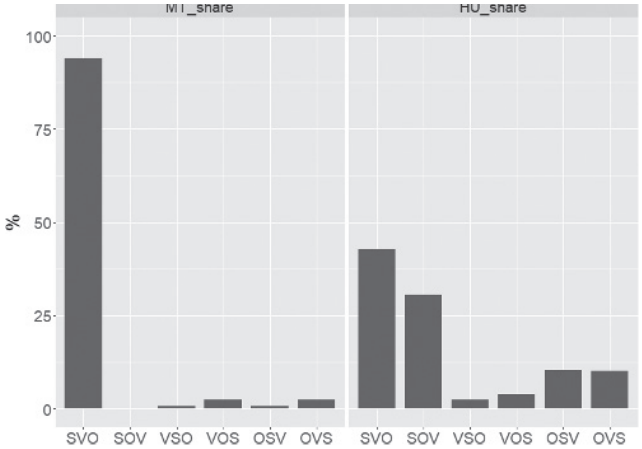


Figure 5: Maltese vs Hungarian – a Greenbergian comparison

A Dryerian analysis provides a more complicated picture (Figure 6):

The primary takeaway here is that both Maltese and Hungarian could safely be classified as SV languages. This, however, does not mean that they behave identically: as we've seen above, no dominant order can be established for Maltese *acl* clauses (and the distribution of the two configurations is almost equal),

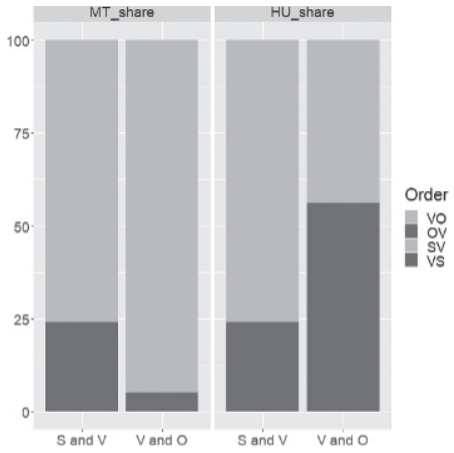


Figure 6: Maltese vs Hungarian – a Dryerian comparison

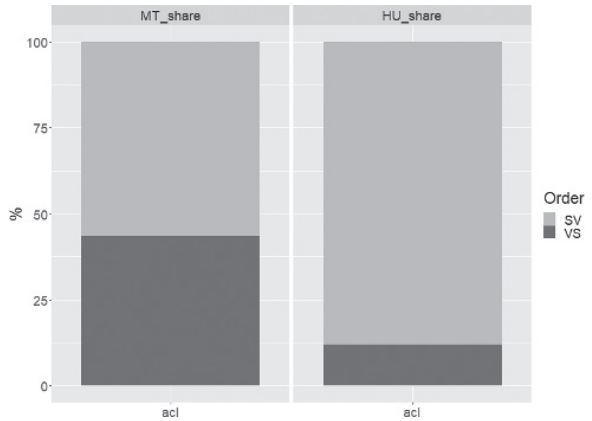


Figure 7: Maltese vs Hungarian – acl clauses

whereas in Hungarian, the share of the VS configuration in *acl* clauses is only half the share of the VS configuration across the board (Figure 7).

As such, Maltese *acl* clauses – and only these clauses – are much more flexible in their ordering of subject and predicate than all other clause types in Maltese; and, conversely, Hungarian *acl* clauses are much more rigid in their ordering of S and V than all other clause types in Hungarian. Whatever this means for the syntax of each respective language, the conclusion one must reach is that despite the numerical similarity, the two languages are actually not that similar in the distribution of the SV and VS constituent order configurations.

A plot of the distribution of VO/OV configurations in both languages (Figure 6) tells a much simpler story. The sharp difference between Maltese (as represented in MUDT v2.7) and Hungarian (as represented in UD v2.5) once again clearly shows that the two languages are not even similar, let alone the same, when it comes to their constituent order. Furthermore, while the data for Hungarian shows that Hungarian (as represented in UD v2.5) cannot be classified as either a VO or an OV language,

it also conforms to the naive expectation regarding constituent order variation in discourse-configurational languages expressed above: the roughly 50-50 distribution of VO and OV is what one would expect if the position of the object were only determined by pragmatic (or, more specifically, information structure) considerations: with only two options (the object is either a topic or it is not), the distribution of VO and OV really should be 1:1.

One might argue that this little comparison does not prove very much: for one, both treebanks are relatively small and thus hardly representative of the language as a whole, especially seeing as the Hungarian UD v2.5 treebank only includes journalistic texts (Zeman, Nivre, Abrams et al. 2019). Additionally, Fabri (2010: 793) may very well be correct in arguing that spoken Maltese is different from written Maltese when it comes to constituent order and so a treebank consisting of spoken materials only might offer a different picture.

As a rebuttal to the second objection, I offer this back-of-the-envelope calculation: MUDT v2.7 contains 1924 clauses featuring a *obj* or a *obl:arg*, of which 94 are OV, for a rate of 5%; the rate of OV in the Hungarian UD v2.5 treebank is 53.1%. If one were to increase the number of OV clauses in MUDT v2.7 five-fold, thus raising the total count of OV clauses to 500 (rounding up), the overall OV share in MUDT v2.7 would climb to only 21% and it would still not even approach the level of OV in the Hungarian UD v2.5 treebank. It would therefore seem more likely that MUDT v2.7 represents this particular aspect of Maltese as a whole rather faithfully (in other words, spoken Maltese may very well be different from written Maltese, but it surely isn't that different), and that this difference between the two treebanks really does represent a real difference between the two languages.

And, to answer the first objection, the composition of the Hungarian v2.5 treebank only underscores this: journalistic texts are typically written in a dry and formal style driven by desire

for clarity and brevity and produced under time crunch, which encourages the use of canned constructions (”journalese”, Suter 1993: 63-68). The fact that even when compared to a relatively balanced MUDT v2.7, the Hungarian UD v2.5 treebank is so different when it comes to the distribution of VO and OV configurations then cannot be explained away by sampling issues. This is doubly true in light of the fact that – as evident from Table 5 below – if one were to compare journalistic texts only, the difference would be even more pronounced: in those types of texts in MUDT v2.7, the share of the OV configuration (3.37%) is even lower than the average in MUDT v2.7 (5%).

Order	**newspaper**	**quasi-spoken**	**fiction**	**non-fiction**
SV	74.03%	68.57%	78.22%	85.41%
VS	25.97%	31.43%	21.78%	14.59%
VO	96.63%	94.12%	93.70%	95.57%
OV	3.37%	5.88%	6.30%	4.43%

Table 5: Constituent order in MUDT v2.7 – Dryerian analysis by genre

Consequently, there are two conclusions to be drawn here: first, Maltese (at least as represented in MUDT v2.7) really is fundamentally different from Hungarian (as represented in the Hungarian UD v2.5 treebank) when it comes to the distribution of constituent order configurations and ipso facto, the two languages cannot belong to the same class defined by a shared property related to constituent order. If one chooses to describe Hungarian as a discourse-configurational language based on the description of its constituent order, it does not seem appropriate to do the same for Maltese. By extension, neither does applying the label “topic-prominent”.

The second conclusion to be drawn from the calculations above is essentially the same as the first one, except broader and methodological rather than descriptive: Borg and Fabri (2016) use

the label "discourse-configurational" as a typological one which is itself somewhat problematic. The real problem, however, is that they do so without considering the entire theory it is based on.[10] As a part of a generative framework, discourse-configurationality is inexorably tied to its fundamental theory of sentence production and its complex conceptual apparatus including base generation, movements and functional projections (cf. Kiss 1995b: 9-10). And even if they were to argue that they only borrow the name and the descriptive information structure concepts behind it (as opposed to the theory of sentence generation), Borg and Fabri fail to consider one crucial property of discourse-configurational languages as defined by Kiss (1995b); the empirical distinction between categorical and thetic statements. In Kiss's wider definition, "[a] language is identified as topic-prominent, more precisely, as a discourse configurational language with property A, if it realizes categorical and thetic judgements in different syntactic structures" (Kiss 1995b: 7-8, see also Chapter 2). Their work does not take this into account and this further invalidates their description of Maltese as a discourse-configurational or a topic-prominent language: such a label, after all, only makes sense within the context of the theory.

Ironically, I've shown here that Maltese actually does employ a different syntactic structure for at least one type of thetic judgments, existential clauses, so taking this into account would support Fabri and Borg's description of Maltese as discourse-configurational as defined in the theory. This argument could be used to make a renewed case for this classification. One could, for example, extend the comparison provided here to other languages and consider the plot in Figure 8, produced from UD v2.7 for Maltese (a putative discourse-configurational language), UD v2.5 for Hungarian (a discourse-configurational language, cf. Kiss

10 This is not the case with Fabri (1993: 140) who describes Maltese as a configurational language, citing the exact definition established in generative literature (see section 1.4).

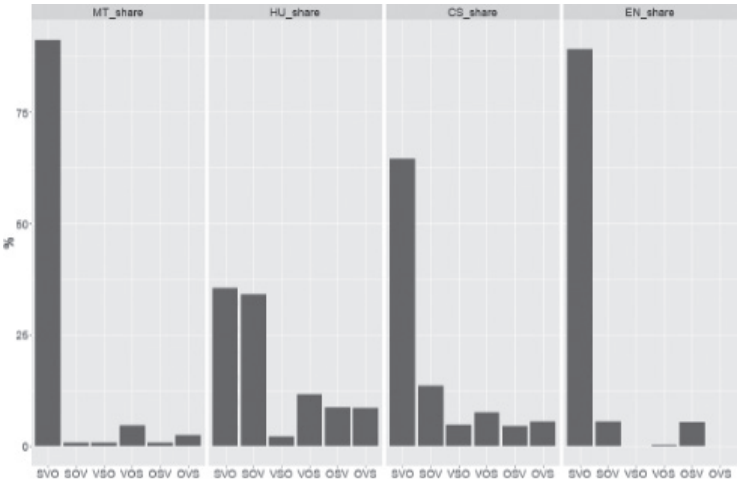


Figure 8: Greenbergian comparison of Maltese, Hungarian, Czech and English

1995b), UD v2.5 for Czech[11] (classified as a language with free or pragmatically determined constituent order, cf. Siewierska and Uhlířová 1998: 109-110) and UD v2.5 for English[12] (a language with a rigid SVO constituent order, cf. Kiss 1995b: 5, 8).

Upon reviewing this data, one could observe that Hungarian and English behave quite differently, as expected from their respective typological classifications. One could also note that Czech is quite different from English and also not that similar to Hungarian. Consequently, one could argue that discourse-configurationality (or indeed topic-orientedness or pragmatical determination of constituent order) is a scale, with Hungarian on one end and English on the other. Whether that would be consistent with the theory is beside the point, what is important is that based on the data above, Maltese (at least as represented in MUDT v2.7) looks much more like a strict SVO language like English, rather than a discourse-configurational language like Hungarian, or a language with pragmatically determined constituent order like Czech.

11 More specifically, the Czech-PDT UD treebank in version 2.5.
12 The English GUM treebank in version 2.5.

Abbreviations

1, 2, 3	first, second, third person	INTJ	interjection
ACC	accusative	M	masculine
COMP	complementizer	NEG	negative
DEF	definite article	PAST	past
EXIST	existentiasl	PL	plural
F	feminine	SG	singular

Acknowledgements

The work underlying this paper was partly financed by grant number APVV-15-0030 from the Slovak Research and Development Agency (APVV), and by the ERC Starting Grant 679083 from the European Research Council, whose support is hereby gratefully acknowledged.

References

Aquilina, Joseph (1959) *The structure of Maltese: A study in mixed grammar and vocabulary*. Malta: The Royal University of Malta.

Borg, Albert & Azzopardi-Alexander, Marie (1997) *Maltese. Lingua descriptive grammars*. London & New York: Routledge.

Borg, Albert & Azzopardi-Alexander, Marie (2009) Topicalisation in Maltese. In: Comrie, Bernard, Fabri, Ray, Hume, Elizabeth, Mifsud, Manwel, Stolz, Thomas & Vanhove, Martine (eds.), *Introducing Maltese linguistics*. Amsterdam: John Benjamins, 71–81.

Borg, Albert & Fabri, Ray (2016) Semantic functions of complementizers in Maltese. In: Boye, Kasper & Kehayov, Peter (eds.), *Complementizer semantics in European languages*. Berlin & Boston: De Gruyter Mouton, 413–447.

Broekhuis, Hans (2008) Derivations and evaluations: Object shift in the Germanic languages. *Studies in Generative Grammar* 97. Berlin & New York: Mouton de Gruyter.

Carnie, Andrew (2013) *Syntax: A generative introduction*. Chichester: Wiley-Blackwell.

Čeplö, Slavomír (2014) An overview of object reduplication in Maltese. In: Borg, Albert, Caruana, Sandro & Vella, Alexandra (eds.), *Perspectives on Maltese linguistics*. Berlin: Akademie Verlag, 201–223.

Čéplö, Slavomír (2018) *Constituent order in Maltese: A quantitative analysis*. Prague: Charles University dissertation.

Choe, Hyoon Sook (1995) Focus and topic movement in Korean and licensing. In: Kiss, Katalin É. (ed.), *Discourse Configurational Languages*. Oxford: Oxford University Press, 269–334.

Chomsky, Noam (1965) *Aspects of the theory of syntax*. Cambridge, Mass: MIT Press.

Corver, Norbert & van Riemsdijk, Henk (1994a) Introduction: approaches to and properties of scrambling. In: Corver, Norbert & van Riemsdijk, Henk (eds.), *Studies on scrambling: movement and non-movement approaches to free word-order phenomena*. Berlin & New York: Mouton de Gruyter, 1–13.

Corver, Norbert & van Riemsdijk, Henk (1994b) *Studies on scrambling: Movement and non-movement approaches to free word-order phenomenon. Studies in generative grammar*. Berlin & New York: Mouton de Gruyter.

de Soldanis, Giovanni Pietro Francesco Agius (1750) *Della lingua punica presentemente usata da maltesi*. Rome.

Derbyshire, Desmond C. (1977) Word order universals and the existence of OVS languages. In: *Linguistic Inquiry* 8.3, 590–599.

Dixon, Robert M. W. (2009) *Basic linguistic theory- vol 1: Methodology*. Oxford: Oxford University Press.

Dryer, Matthew S. & Haspelmath, Martin (eds.) (2013) *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info (checked 02/29/2020)

Dryer, Matthew S. (1989a) Discourse-governed word order and word order typology. In: *Belgian journal of linguistics* 4, 69–90.

Dryer, Matthew S. (1989b) Large linguistic areas and language sampling. In: *Studies in Language* 13, 257–292.

Dryer, Matthew S. (1992) The Greenbergian word order correlations. In: *Language* 68.1(1), 81–138.

Dryer, Matthew S. (1997) On the 6-way word order typology. In: *Studies in Language*, 69–103.

Dryer, Matthew S. (2007) Word order. In: Shopen, Timothy (ed.), *Language typology and syntactic description. Second edition- vol. I: Clause Structure.* Cambridge: Cambridge University Press, 61–131.

Dryer, Matthew S. (2013a) Determining dominant word order. In: Dryer, Matthew S. & Haspelmath, Martin (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, n/a.

Dryer, Matthew S. (2013b) On the six-way word order typology, again. In: *Studies in Language* 37.2, 267–301.

Dryer, Matthew S. (2013c) Order of subject, object and verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Fabri, Ray (1993) *Kongruenz und die Grammatik des Maltesischen* (Agreement and the Grammar of Maltese). Tübingen: Niemeyer.

Fabri, Ray (2010) Maltese. In: *Revue belge de philologie et d'histoire* 88.3, 791–816.

Fabri, Ray & Borg, Albert (2002) Topic, focus and word order in Maltese. In: Abderrahim, Youssi, Benjelloun, Fouzia, Dahbi, Mohamed & Iraqui-Sinaceur, Zakia (eds.), *Aspects of the dialects of Arabic today. Proceedings of the 4th Conference of*

the International Arabic Dialectology Association (AIDA). Marrakesh, Apr. 1-4.2000. In Honour of Professor David Cohen. Rabat: Amapatril, 354–363.

Fabri, Ray & Borg, Albert (2017) Modifiers and complements within the Maltese verb sequence. In: Saade, Benjamin & Tosco, Mauro (eds.), *Advances in Maltese linguistics*. Berlin & New York: De Gruyter Mouton, 67–86.

Féry, Caroline & Shinichiro Ishihara (2016a) Introduction. In: Féry, Caroline & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*. Oxford: Oxford University Press, 3–15.

Féry, Caroline & Shinichiro Ishiharaeds (eds.) (2016b) *The Oxford handbook of information structure*. Oxford: Oxford University Press.

Firbas, Jan (1964) On defining the theme in functional sentence analysis. In: *Travaux Linguistiques de Prague* 1, 267–280.

Firbas, Jan (1992) *Functional sentence perspective in written and spoken communication*. Cambridge: Cambridge University Press.

Gatt, Albert & Čéplö, Slavomír (2013) Digital corpora and other electronic resources for Maltese. *Talk at the Corpus Linguistics 2013 conference at Lancaster University*.

Givón, Talmy (2001) *Syntax. An introduction- vol. II*. Amsterdam & Philadelphia: John Benjamins.

Greenberg, Joseph H. (1966) Some universals of grammar with particular reference to the order of meaningful elements. In: Greenberg, Joseph H. (ed.), *Universals of language*. Mass: MIT Press, 73–113.

Hale, Ken (1983) Warlpiri and the grammar of non-configurational languages. In: *Natural Language & Linguistic Theory* 1.1, 5–47.

Harris, Randy Allen (1995) *The linguistics wars*. Oxford: Oxford University Press.

Itkonen, Esa (2005) Concerning the synthesis between intuition-based study of norms and observation-based study of corpora. In: *SKY Journal of Linguistics* 18, 357–377.

Kalmár, Ivan & Agius, Dionisius (1981) Verb-subject order in Maltese. In: *Journal of Maltese Studies* 14, 22–31.

Kiss, Katalin É. (ed.) (1995a) *Discourse configurational languages*. Oxford: Oxford University Press.

Kiss, Katalin É. (1995b) Discourse configurational languages: Introduction. In: Kiss, Katalin É. (ed.), *Discourse configurational languages*. Oxford: Oxford University Press, 3–27.

Krause, Thomas & Zeldes, Amir (2016) ANNIS3: A new architecture for generic corpus query and visualization. In: *Digital Scholarship in the Humanities* 31.1, 118–139.

Krier, Fernande (1976) *Le maltais au contact de l'italien. Etude phonologique, grammatical et sémantique*. Hamburg: Helmut Buske Verlag.

Krifka, Manfred & Musan, Renate (2012) Information structure: Overview and linguistic issues. In: Krifka, Manfred & Musan, Renate (ed.), *The Expression of Information Structure*. Berlin/Boston: De Gruyter Mouton, 1–44.

Lewis, Paul M., Simons Gary F. & Fennig, Charles D. (2016) *Ethnologue: Languages of the world*, Nineteenth edition. Dallas, Texas: SIL International.

Marty, Anton (1897) Über die Scheidung von grammatischem, logischem und psychologischem Subjekt resp. Prädikat. In: *Archiv für Systematische Philosophie* 3, 174–189.

Marty, Anton (1895) Über Subjektlose Sätze und das Verhaltniss der Grammatik zu Logik und Psychologie [Sechster Artikel]. In: *Vierteljahrsschrift für wissenschaftliche Philosophie* 19, 19–87, 263–334.

Mathesius, Vilém (1939) O tak zvaném aktuálním členění věty. In: *Slovo a slovesnost* 5, 171–174.

Mathesius, Vilém (1961) Obsahový rozbor současné angličtiny na základě obecně lingvistickém. Praha: Nakladatelství Československé akademie věd.

McEnery, Tony & Hardie, Andrew (2011) *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.

Milička, Jiří (2014) Menzerath's Law: The whole is greater than the sum of its parts. In: *Journal of Quantitative Linguistics* 21(2), 85–99.

Payne, Thomas E. (1997) *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.

Puskás, Genoveva (2000) *Word order in Hungarian. The syntax of Ā-positions*. Amsterdam: John Benjamins.

Ross, John R. (1967) *Constraints on variables in syntax*. PhD thesis. Massachusetts Institute of Technology.

Siewierska, Anna & Uhlířová, Ludmila (1998) An overview of word order in Slavic languages. In: Siewierska, Anna (ed.), *Eurotyp: Typology of languages in Europe- vol 1: Constituent order in the languages of Europe. Empirical approaches to language typology* 20.1. Berlin & New York: Mouton de Gruyter.

Song, Jae Jung (ed.) (2011a) *The Oxford handbook of linguistic typology*. Oxford: Oxford University Press.

Sutcliffe, Edmund (1936) *A grammar of the Maltese language. With chrestomathy and vocabulary*. Oxford: Oxford University Press.

Suter, Hans-Jürg (1993). *The wedding report. A prototypical approach to the study of traditional text types*. Amsterdam: John Benjamins.

Vella, Francis (1831) *Maltese grammar for the use of the English*. Leghorn: Glaucus Masi.

Vella, Joseph (1970) *A comparative study in Maltese and Libyan (Benghazi Dialect). Phonetics, morphology, syntax & lexicon*. Malta: The Royal University of Malta (Unpublished PhD thesis).

von der Gabelentz, Georg (1869) Ideen zu einer vergleichenden Syntax. In: *Zeitschrift für Völkerpsychologie und Sprachwissenschaft* 6, 376–384.

Weil, Henri (1844) *De l'ordre des mots dans les langues anciennes comparées aux langues modernes*. Paris: Joubert.

Zeman, Daniel, Nivre, Joakim, Abrams, Mitchell et al. (2019) *Universal dependencies 2.5*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11234/1-3105

Zeman, Daniel, Nivre, Joakim, Abrams, Mitchell et al. (2020). *Universal dependencies 2.7*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11234/1-3424

Zubizarreta, Maria Luisa (1998). *Prosody, focus, and word order*. Cambridge, Massachusetts & London, England: The MIT Press.

NEGATION IN MALTESE

Christopher Lucas

Abstract

This article gives an overview of the key facts relating to the expression of negation in contemporary Standard Maltese. Issues considered include: anaphoric negation, constituent negation, sentential negation expressed both with the particle *mhux* and the bipartite construction *ma…-x*, factors governing omission of one or both elements of the bipartite construction, the interaction of negation with indefinite pronouns, and the analysis of *-x* in non-negative contexts.

Dan l-artiklu joffri deskrizzjoni tal-fatti ewlenin marbutin man-negazzjoni fil-Malti Standard kontemporanju. Fost it-temi trattati hemm in-negazzjoni anaforika, in-negazzjoni tal-kostitwenti, in-negazzjoni sentenzjali espressa kemm bil-partiċella *mhux* kif ukoll bil-binja bipartita *ma…-x*, il-fatturi li jirregolaw l-ommissjoni ta' komponent wieħed jew taż-żewġ komponenti li jiffurmaw il-binja bipartita, l-interazzjoni tan-negazzjoni mal-pronomi indefiniti, u l-analiżi ta' *-x* f'kuntesti mhux negattivi.

1. Introduction

In formal logic, negation is an operation which is applied to a proposition, and which has the effect of reversing the conditions under which that proposition is true. For example, consider the proposition expressed by the English sentence *I am taller than you*. The truth of this proposition depends on certain facts about the world at the time the sentence is uttered: most importantly, that the speaker's height is indeed greater than that of the addressee. If these conditions hold, then this sentence expresses a true proposition, while the negative version of this sentence – *I am not taller than you* – expresses a false proposition. Thus, whenever a proposition *p* is true, the negation of that proposition (symbolized as ¬*p* in formal logic) is false, and vice versa.

As far as it is possible to tell, all human languages have some means of modifying utterances in ways similar to the negation operation of formal logic. The present article sets out some of the most important ways in which this semantic notion of negation manifests itself in Maltese.[1] We consider anaphoric negation (§2), sentential negation (§3), constituent negation (§4), negation and indefinite pronouns (§5), and non-negative uses of the suffix *-x* (§6).

2. Anaphoric negation

Some languages, including Maltese and English, have a specialized morpheme, distinct from the main sentence negator(s), to succinctly deny the truth of a salient proposition from the immediately preceding discourse. In Maltese and English these are *le* and *no*, respectively (the corresponding affirmative forms being *iva* and *yes*). Maltese *le* is cognate with Classical Arabic

1 The use of the term Maltese in this article should be understood as referring in all cases to contemporary Standard Maltese, unless otherwise indicated.

lā, but unlike the latter it only functions as an anaphoric negator in the language, as in (1). English *no* has an additional function as a negative determiner (as in *no books*) not shared by Maltese *le*, which expresses this meaning with a dedicated item *ebda*. The interaction of indefinite pronouns of this sort with negation in Maltese is addressed in §5.

(1) Korpus Malti v3.0 news148902

Kif	*tista'*	*tgħid*	*li*	*xi ħadd*	*huwa*
how	MOD.IPFV.2SG	say.IPFV.2SG	COMP	someone	3SGM

terrorist	*u*	*ieħor*	***le****?*
terrorist	CONJ	other	no

'How can you say that someone is a terrorist and someone else is not?' (Lit.: '...and someone else, no')

3. Sentential negation

Apart from anaphoric negation, what we might think of as the most basic type of negation in any language is sentential negation. This term refers to the means used to render a sentence negative in the sense outlined in §1. Many languages have several different morphemes or constructions to express sentential negation, depending on the precise context or function, and Maltese is no exception. The different ways in which sentential negation is expressed in Maltese are addressed in the following subsections. Negation that operates below the level of the sentence is dealt with in §4.

3.1 *Standard negation*

Standard negation is a term coined by Payne (1985). It refers to "the non-emphatic negation of a lexical main verb in a declarative main clause" (van der Auwera & Krasnoukhova 2020: 91). Standard negation in Maltese is expressed by a bipartite construction *ma...-x*, as in (2).

(2) Korpus Malti v3.0 news139511

...ħafna	*nies*	***ma***	*jifhmu-**x***	*x'-inkunu*
much	people	NEG	understand.IPFV.3PL-NEG	what-AUX.IPFV.1PL
rridu		*ngħidu.*		
want.IPFV.1PL		say.IPFV.1PL		

'Many people don't understand what we are trying to say.'

This bipartite construction is shared with a number of Arabic dialects, especially those of North Africa to which Maltese is most closely related (see Lucas 2018, Lucas & Alluhaybi 2022 for details). The second element of the construction, *-x*, like the indefinite pronoun *xejn* and the particle *xi* (discussed in §5), ultimately derives from the Arabic word *šayʔ* 'thing'. The common historical process whereby a more lexical item (such as *šayʔ*) comes to be recruited for a more grammatical function (often in a highly phonologically reduced form, as with *-x*) has been known since at least Meillet (1912) as grammaticalization. The particular type of grammaticalization evidenced by the development of Arabic *šayʔ* into the Maltese negative element *-x* has been known since Dahl (1979) as Jespersen's cycle, after the Danish linguist Otto Jespersen, who, in a (1917) work, was one of the first authors to describe this cycle as it occurred in the history of English and French. For more details on Jespersen's cycle in the languages of Europe and the Mediterranean, see Willis et al. (2013a) and Breitbarth et al. (2020).

3.2 Bipartite negation in other contexts

In addition to regular main verbs, the bipartite *ma...-x* construction is also typically used to negate the copula in most of its various forms, as well as so-called pseudo-verbs, modal and other auxiliary verbs, and verbs in subordinate clauses. We consider each of these in turn in the following.

3.2.1 *Copular negation*

As Stassen (1996) and Camilleri & Sadler (2019) among others describe, Maltese has a number of different constructions that function as equivalents of the English 'to be' copula verb. One of these, *qiegħed*, is (from an etymological point of view) a participle and is addressed in §3.3.2. Two of them, *kien* and *jinsab*, are verbs which, while exhibiting various irregularities, are negated just like any other verb, as illustrated in (3).

(3)

a. Korpus Malti v3.0 news11011

F'-Mater Dei	*ġie*	*ċċertifikat*	*li*
in-PN	come.PFV.3SGM	certify.PTCP.PASS	COMP

ħajt-u	***ma***	***tinsab-x***	*f-il-periklu.*
life-3SGM	NEG	COP.IPFV.3SGF-NEG	IN-DEF-danger

'In Mater Dei hospital it has been confirmed that his life is not in danger.'

b. Korpus Malti v3.0 academic189

Il-kronoloġiji	*ta-l-paġni*	*mħassra*
DEF-chronology.PL	GEN-DEF-PAGE.PL	damage.PTCP.PASS.PL

f'-dak	*il-perjodu*	***ma***	***kinu-x***	*disponibbli.*
in-DEM	DEF-period	NEG	COP.PST.3PL-NEG	available

'The chronologies of the pages damaged during that period were not available.'

More interesting from the point of view of negation are the other two copular constructions found in Maltese: one involving a third-person pronominal copula *hu(wa)/hi(ja)/huma*, and one with no copula. In the negative, these two constructions collapse into one, in which the copula takes the form of any of the personal pronouns circumfixed with *ma...-x*, as shown in (4) and Table 1.[2]

2 The first person singular negative pronominal copula also has the alternative forms *m'jiniex*, *ma jienx* and *m'jienx*. The second person singular has the alternative form *m'intx*. All persons and numbers additionally appear without suffixed *-x* in the contexts set out in §3.3.1.

(4)

a. Korpus Malti v3.0 news72278

Aħna	***m'-aħnie-x***	*hawn*	*biex*	*nagħmlu*	*l-gwerer...*
1PL	NEG-1PL-NEG	here	PURP	do.IPFV.1PL	DEF-war.PL

'We are not here to fight wars...'

b. Korpus Malti v3.0 news190513

Dan	***m'huwie-x***	*xogħol*	*faċli,*	*jieħu*	*żmien...*
DEM	NEG-3SGM-NEG	work	easy	take.IPFV.3SGM	time

'This is no easy task, it takes time...'

Person	**Singular**	**Plural**
1st	m'iniex	m'aħniex
2nd	m'intix	m'intomx
3rd masc.	mhuwiex/mhux	mhumiex
3rd fem.	mhijiex/mhix	

Table 1: Paradigm of the Maltese negative pronominal copula

A very common alternative to full person–number–gender agreement of the negative copula as seen in (4) is the use of *mhux* as a frozen form in negative copular sentences with subjects of any person/number/gender, as in (5).

(5)

a. Korpus Malti v3.0 culture1149

Din	***mhux***	*l-ewwel*	*darba*	*li*	*Ira*	*tkellmet*	*kontra*
DEM.F	NEG	DEF-first	time	COMP	PN	speak.PFV.3SGF	against

l-bullying...

DEF-bullying

'This is not the first time that Ira has spoken out against bullying...'

b. Korpus Malti v3.0 parl12135

intom	***mhux***	*biss*	*union,*	*imma*	*assoċjazzjoni...*
2PL	NEG	only	union	but	association

'you are not just a union, but also an association...'

This frozen element *mhux* is in fact the basic negator for various constructions that in contemporary Maltese cannot be seen as copular, as discussed in §3.3.2 and §4. This form, as well as the fully

inflected forms shown in Table 1, are, however, also sometimes used for the negation of ordinary verbs, in contexts where bipartite *ma...-x* would be the typical negative construction. As shown by Al-Sayyed & Wilmsen (2017), use of *mhux* as a negator of ordinary verbs is associated with various pragmatic or metalinguistic functions, typically centred around the denial of a discourse-active presupposition (cf. Schwenter 2005; Hansen 2013 for discussion of similar phenomena in other languages), while Spagnol (2009) argues that use of *mhux* with imperfective verbs triggers a progressive interpretation of the verb. Both these properties are illustrated in (6), with (6b) demonstrating that they also hold true for the fully inflected forms of the negative copula used as a verbal negator.

(6)

a. Korpus Malti v3.0 parl3502 (Al-Sayyed & Wilmsen 2017: 163)

mhux	*nistaqsi*	*iżda*	*nittallab*	*lil-l-ministeri*
NEG	ask.IPFV.1SG	but	beg.IPFV.1SG	OBJ-DEF-ministry.PL
konċernati...				
concern.PTCP.PASS.PL				

'I am not asking but begging from the ministries concerned...'

b. Korpus Malti v3.0 news127417

M-inie-x	*nikteb*	*dan*	*l-artiklu*	*biex*	
NEG-1SG-NEG	write.IPFV.1SG	DEM	DEF-article	PURP	
nattakka	*lil*	*xi_ħadd,*	*imma...*		
attack.IPFV.1SG	OBJ	someone	but		

'I am not writing this article to attack someone, but...'

3.2.2 *Negation of pseudo-verbs*

On the other hand, there is a class of predicates in Maltese which are etymologically non-verbal, but which exhibit various verb-like properties, including in a number of cases being typically negated with the bipartite *ma...-x* construction, as illustrated in (7). This class of predicates, which includes items such as existential *hemm*, possessive *għand-*, and *għad-* 'still', are usually given the collective label of pseudo-verbs, and the fact that they are negated in the same

way as verbs is treated by Comrie (1982) as a key piece of evidence in favour of analysing these items as (irregular) verbs.[3] See Peterson (2009) for a detailed discussion of Maltese pseudo-verbs.

(7)

a. Korpus Malti v3.0 academic8

Aħna	***m'-ghand-nie-x***	*involviment*	*dirett*	*f-il-linja*
1PL	NEG-POSS-1PL-NEG	involvement	direct	in-DEF-line

ta-l-produzzjoni.
GEN-DEF-production
'We do not have direct involvement in the production line.'

b. Korpus Malti v3.0 academic10

M'-hemm-x	*dubju*	*li*	*s-suq*	*globali*	*se*
NEG-EXS-NEG	doubt	COMP	DEF-market	global	FUT

jkompli	*jeżisti...*
continue.IPFV.3SGM	exist.IPFV.3SGM

'There is no doubt that the global market will continue to exist...'

3.2.3 *Negation of auxiliary verbs*

A salient feature of Maltese syntax is the prevalence of extended chains of finite verbs, as in (8) (cf. Stolz 2009; Fabri & Borg 2017).

(8) BCv3: 1993 Immanuel Mifsud - Il-Ktieb tas-Sibt Filgħaxija (Čéplö 2018: 145)

Issa	*se*	*jkoll-i*	*nerġa'*	*nibda*	*nistenna*
NOW	FUT	MOD-1SG	return.IPFV.1SG	begin.IPFV.1SG	wait.IPFV.1SG

'Now I will have to once more start waiting.'

Negation interacts with such verbal chains in interesting and complex ways, a thorough examination of which remains a desideratum for future research. Here let us simply make a few key observations that arise from the fact that, in theory, any of the verbs in such a chain should be able to host negation.

3 Note that clearly non-verbal predicates (such as nominal, adjectival or prepositional phrases) in copular clauses cannot be negated this way in Maltese, as illustrated in (i):

(i) **Il-pop_music* *m'-interessanti-x.*
DEF-pop_music NEG-interesting-NEG
Intended meaning: 'Pop music is not interesting.'
[Adapted from Korpus Malti v3.0 parl18: *l-pop music mhux interessanti*]

First, note that more than one verb in such a chain can host negation, as pointed out by Stolz (2009: 153), in which case we have logical double negation, with two negatives cancelling each other out to form an affirmative, as illustrated in (9).

(9) Korpus Malti v3.0 news126780 (cf. Stolz 2009: 153)

Ma	*nistgħu-**x***	***ma***	*naħsbu-**x***	*kif...*
NEG	can.IPFV.1PL-NEG	NEG	think.IPFV.1PL-NEG	how

'We cannot not consider how...'

The example in (9) involves a sequence of two negated verbs. This does not seem to represent the upper limit from a syntactic point of view. But the one example in the 250-million-word Korpus Malti v3.0 of three successive negated verbs in a single clause, shown in (10), appears to be a case of what has come to be known as misnegation or overnegation:[4] a frequent phenomenon distinct from that of negative concord (discussed in §5), whereby speakers become confused as to the number of negations required to convey their meaning, such that the literal meaning of an utterance is the opposite of what is intended (and usually understood without hesitation by the addressee).

(10) Korpus Malti v3.0 news60485

Iżda	***ma***	*nista-**x***	***ma***	*nerġa-**x***
but	NEG	MOD.IPFV.1SG-NEG	NEG	return.IPFV.1SG-NEG
ma	*nirringrazzja-**x** [...]*		*lil-l-president...*	
NEG	thank.IPFV.1SG-NEG		OBJ-DEF-president	

Literal meaning: 'But I cannot fail to again not thank the president...'

Intended meaning: 'But I can't not thank the president once more...'

However, the norm is just one negated verb per chain, with the choice of which verb is negated determined, as in (9), by which predicate the speaker wishes to deny holds. This is easiest to see when one of the verbs has a modal value (i.e. meanings such as 'can', 'must', 'ought', etc.), as in (11), where

4 These terms appear to have been invented by contributors to the linguistics blog *Language Log*. See https://languagelog.ldc.upenn.edu/nll/?cat=273 (accessed 24/5/2023) for examples.

the relative scope of the modal, the predicate, and the negation is clear.

(11)

a. Korpus Malti parl547

kuntrattur [...]	***jista'***		***ma***	***jkun-x***	
contractor	MOD.IPFV.3SGM		NEG	be.IPFV.3SGM-NEG	
kopert	*b'*	*"ċittadin*	*Malti",*	*imma...*	
COVER.PTCP.PASS	PREP	citizen	Maltese	but	

'It is possible for a contractor **not** to be covered by [the designation] "Maltese citizen", but...'

b. Korpus Malti v3.0 european9802

Riskju	*ġdid [...]*	*jirrekjedi*		*prodott*	*ta-l-assigurazzjoni*	
risk	new	require.IPFV.3SGM		product	GEN-DEF-insurance	
kompletament	*ġdid,*	*u*	***ma***	***jista-x***		
completely	new	CONJ	NEG	MOD.IPFV.3SGM-NEG		
jkun	*kopert*			*b'-żidiet*	*jew*	
be.IPFV.3SGM	COVER.PTCP.PASS			PREP-addition.PL	or	
modifikazzjonijiet	*f'-prodott*			*ta-l-assigurazzjoni*		*eżistenti.*
modification.PL	PREP-product			GEN-DEF-insurance		existing

'A new risk requires a completely new insurance product, and it is not possible for the risk to be covered by additions or modifications to the existing insurance product.'

In cases of verb chains involving the aspectual auxiliary *kien*, which has no conceptual content, it is not usually possible for negation to have scope over the main verb only (and not also *kien*). To see why, consider *it was the case that he didn't laugh* versus *it wasn't the case that he laughed*. Just as there is no difference in the literal meaning of these sentences, in the same way no obvious difference in meaning could be achieved by moving the expression of negation from the auxiliary to the main verb in an example such as (12). In such cases it is the auxiliary that carries negation by default, a fact probably best explained by what Horn (2001: 292) calls the Negative First Principle, following Jespersen's (1933: 297) observation of the tendency "to put the negative word or element as early as possible, so as to leave no doubt in the mind of the hearer as to the purport of what is said."

(12) Korpus Malti v3.0 parl111

Intom	***ma***	***kontu-x***	*titkellmu*	*magħ-hom...*
2PL	NEG	AUX.PST.2PL-NEG	speak.IPFV.2PL	with-3PL

'You didn't use to talk to them...'

Essentially the only exception to this generalization concerns the verb *felaħ* 'to manage to, to be able to afford to; to thrive, be in good health'. In its 'manage' meaning it is unexceptional, as illustrated in (13).

(13) Korpus Malti v3.0 news151867

...għax	***ma***	***kien-x***	*jiflaħ*
because	NEG	AUX.PST.3SGM-NEG	manage.IPFV.3SGM

jirrispondi	*għal-l-mistoqsijiet.*
respond.IPFV.3SGM	PREP-DEF-question.PL

'...because he didn't have the strength to reply to the questions.'

In its 'thrive' meaning, however, this verb is used predominantly in the negative,[5] where it takes on the meaning not of merely not thriving but of being actively unwell. Here we see that, even in combination with *kien*, a distinction of scope becomes relevant, as it was for modal *jista'* above: a speaker may feel it necessary to make clear that she is not referring merely to a lack of vigorous good health, but in fact to the active presence of poor health. This explains why combinations of this verb in this meaning with *kien* typically show negation on *felaħ*, as in (14) (contrast with (12) and (13)).

(14) Korpus Malti v3.0 news132671

...kell-hom	*jinżlu*	*mingħajr*	*il-goalkeeper*
MOD.PST-3PL	descend.IPFV.3PL	without	DEF-goalkeeper

regolari	*tagħ-hom,*	*Nicky Gouder,*	*li*	*kien*	***ma***
regular	GEN-3PL	PN	COMP	AUX.PST.3SGM	NEG

jiflaħ-x.

thrive.IPFV.3SGM-NEG

'...they had to turn out without their regular goalkeeper, Nicky Gouder, who was unwell.'

5 Other verbs used predominantly in the negative include *ħamel* 'to bear, tolerate', and, in the negative imperative (see § 3.3.1), *iwworja* 'to worry'.

3.2.4 *Negation in subordinate clauses*

Unlike in certain languages such as Latin or Greek, there are no particles or constructions in Maltese that are specialized for the negation of predicates in subordinate clauses. Such predicates are negated in exactly the same way as their main-clause counterparts described in the rest of this article. There are also no obvious respects in which the syntax of negation in subordinate clauses in Maltese differs in interesting ways from other better described languages. Nevertheless, it is worth mentioning here an interesting phenomenon in this domain that Maltese shares with many (perhaps all) languages, namely what is called "neg-raising" (see Horn 1989: §5.2 for a detailed pragmatic account of this phenomenon). Neg-raising is the phenomenon whereby, with a restricted class of experiential predicates such as 'think', 'believe', and 'want', negation appears in a higher clause than where it is interpreted.

To illustrate the phenomenon, consider first the ordinary case with non-neg-raising predicates. Here we see that, just as with verb chains in a single clause, in a main-clause–subordinate-clause sequence, the locus of the negative particle(s) is determined by the semantic scope of negation relative to the main-clause and subordinate-clause predicates. This is shown in (15), taken from Borg & Azzopardi-Alexander (1997: 93), where we see that when there is an order not to move, negation attaches to 'move', whereas when it is denied that there was an order to move, negation attaches to 'order'.

(15)

a. *Ordna-l-u* ***ma*** *jiċċaqlaq-**x**.*
order.PFV.3SG-DAT-3SG NEG move.IPFV.3SG-NEG
'He ordered him not to move.'

b. ***M'**-ordna-l-u-**x*** *jiċċaqlaq* *minn* *post-u.*
NEG-order.PFV.3SG-DAT-3SG-NEG move.IPFV.3SG from place-3SGM
'He did not order him to move from his place.'

However, with neg-raising predicates such as *ħaseb* 'think' or *emmen* 'believe', negation is typically marked on the higher (neg-raising) predicate and nevertheless interpreted in the lower clause. Thus, in (16), also taken from Borg & Azzopardi-Alexander (1997: 93), the speaker is not denying that there is something she has a belief about; she is stating her belief, or worry, that she cannot afford the sum in question.

(16) ***Ma*** *naħsib-**x*** *li* *niflaħ* *inħallas*
NEG think.IPFV.1SG-NEG COMP afford.IPFV.1SG spend.IPFV.1SG
daqshekk *f-ix-xahar*.
such in-DEF-month
'I don't think I can afford to pay so much every month.'

3.3 Single negation

3.3.1 *Omission of either ma or -x*

There are a number of contexts in which one or even both of the two elements of the bipartite construction are omitted. The second element, *-x*, is omitted: i) when an element closely related to the negated predicate is an indefinite pronoun or adverb, as in (17) (see §5 for further details); and ii) in co-ordinated negative sentences involving the focus particle *lanqas*, as in (18), where we also observe a unique aspect of this construction, namely the use of *la* in place of *ma* as the negator of the first element.[6]

6 There is also a distinct use of *lanqas* as a negative scalar focus particle 'not even', where it behaves similarly to indefinite pronouns in that it generally triggers omission of *-x*, as in (i). See Čéplö & Lucas (2020) for more details.
(i) Korpus Malti v3.0 literature21
U ***lanqas*** *jiena* **ma** *stajt* *norqod.*
CONJ FOC 1SG NEG can.PFV.1SG sleep.IPFV.1SG
'And not even I could sleep.'

(17) Korpus Malti v3.0 academic12

Il-ħajja	*f'-dawn*	*l-istituti*		***qatt***	***ma***
DEF-life	in-DEM.PL	DEF-institution.PL		never	NEG
kienet	*waħda*	*ta'*	*lussu.*		
COP.PST.3SGF	one.F	POSS	luxury		

'The life in those institutions was never one of luxury.'

(18) Korpus Malti v3.0 academic12

La	*kienu*	*kapaċi*	*jaqraw*	*u*	***lanqas***
NEG	COP.PST.3PL	capable.PL	read.IPFV.3PL	CONJ	FOC
jiktbu...					
write.IPFV.3PL					

'They were able neither to read nor to write...'

The first element of the bipartite construction, *ma*, is omitted with negative imperatives, as in (19).[7]

(19) Korpus Malti v3.0 religion458

*Tarmi-**x***	*barra,*	*tħammiġ-**x**,*	*ibża'*
throw.IPFV.2SG-NEG	outside	dirty.IPFV.2SG-NEG	fear.IMP.2SG
għall-ambjent...			
on.DEF-environment			

'Don't litter, don't make a mess, look after the environment...'

When we combine negative imperatives with indefinite pronouns, both the first and the second elements of the *ma...-x* construction are omitted, as in (20).

(20) Notice observed in Valletta

Tarmi	***xejn***	*hawn.*
throw.IPFV.2SG	nothing	here

'Don't throw anything here.'

7 Historically speaking it is probably not correct to say that it is *ma* that is omitted from negative imperative sentences, since in older Maltese texts we find *la* as the preverbal negator in these, as illustrated in (i):

(i) Traditional Maltese song

Ninni	*la*	*tibki-**x***	*iżjed.*
sleep.IMP.2SG	NEG	cry.IPFV.2SG-NEG	more

'Go to sleep, don't cry any more.'

3.3.2 *Single negation with mhux*

We saw in §3.2 that the basic rule is that the bipartite *ma...-x* construction is reserved for verbs. If we wish to make this rule exceptionless, then we need to analyse the negative pronominal copula (§3.2.1) and a number of pseudo-verbs (§3.2.2) as irregular kinds of verbs, and, as we have seen, there are researchers who have made such arguments. Another kind of predicate with verbal qualities, which is, however, not negated with *ma... -x* but with *mhux*, is the participle, as illustrated in (21) for active and passive participles respectively.

(21)

a. Korpus Malti v3.0 news109720

L-argument	*tagħ-hom*	***mhux***	***nieżel***	*tajjeb*
DEF-argument	GEN-3PL	NEG	descend.PTCP.ACT	well

din	*id-darba.*
DEM.F	DEF-time

'Their argument is not really sound on this occasion.'

b. Korpus Malti v3.0 parl453

...għand-ek	*issib*	*kriterju*	*ieħor*	*li*	***mhux***
MOD-2SG	find.IPFV.2SG	criterion	other	COMP	NEG

marbut	*ma-l-eżami*	*ta-l-mezzi.*
link.PTCP.PASS	PREP-DEF-test	GEN-DEF-wealth.PL

'...you need to find another criterion that is not connected to means testing.'

Participles are, by definition, nominal or adjectival elements derived from verbs. As such, we could analyse the participle-containing clauses in (21) simply as (non-verbal) copular clauses, and explain the use of *mhux* rather than *ma...-x* here in this way (cf. §3.2.1). However, it seems that at an early stage in the (pre-) history of Maltese (as in apparently all Arabic varieties) *mhux* (or its equivalents in Arabic varieties) was felt to be an appropriate negator for participles specifically, no matter how verb-like their function. This would then explain why verb phrases containing aspectual particles such as future-marking *se* and progressive-

marking *qed*, which derive historically from the participial forms *sāyir 'going' and *qāʕid 'sitting' (> Maltese *qiegħed*), are negated with *mhux*, rather than with *ma...-x* as one might otherwise have expected:[8]

(22)

a. Korpus Malti v3.0 academic12

...imma	*żgur*	*li*	*l-Maltin*	***mhux***	***se***	***jsibu***
but	sure	COMP	DEF-Maltese.PL	NEG	FUT	find.IPFV.3PL
post	*aħjar*	*minn*	*Melbourne.*			
place	better	PREP	PN			

'...but it is certain that the Maltese will not find a better place than Melbourne.'

b. Korpus Malti v3.0 culture2700

Jekk	*il-karozza*	*tinduna*	*li*	*s-sewwieq*	***mhux***
COMP	DEF-car	notice.IPFV.3SGF	COMP	DEF-driver	NEG
qed	***ihares***...				
PROG	look.IPFV.3SGM				

'If the car notices [through sensors and cameras] that the driver is not paying attention...'

Similarly, the full form *qiegħed*, from which *qed* derives, retains the participial type of negation with *mhux* in its present-day function as a copula, as illustrated in (23).

(23) Korpus Malti v3.0 opinion1717

Imma	*bħalissa*	***mhux***	*qiegħed*	*l-isptar.*
but	currently	NEG	COP	DEF-hospital

'But he is not in hospital at the moment.'

8 Note, however, that in a number of Maltese dialects *qed* (or even the full form *qiegħed*) is treated as a pseudo-verb (cf. § 3.2.2) and negated with *ma…-x* instead of mhux:

(i) Dialect of Mġarr (Vanhove 1993: 131)

ma-ʔet-š	*naḥdim*	*[≈ Ma qedx naħdem.]*
NEG-PROG-NEG	work.IPFV.1SG	

'I am not working.'

4 Constituent negation

Klima (1964) introduced a distinction between sentential negation and constituent negation. Constituent negation is sub-sentential: it is when negation has scope over some word or phrase that is only a part of a sentence or clause. This therefore includes negative prefixes, such as the *in-* prefix in Maltese words of Italian origin, as in *in-ċertezza* 'un-certainty' or *in-direttament* 'in-directly'. But it also includes phrases made negative. This is illustrated in (24), where we see that the same negator *mhux* that is used for negating non-verbal predicates, participles, and prefixes derived from participles is also used for constituent negation.

(24) Korpus Malti v3.0 news85703

...għand-i	*idea*	***mhux***	*ħażin*	*ta'*	*dawn*	*l-affarijiet...*
POSS-1SG	idea	NEG	bad	GEN	DEM.PL	DEF-matter.PL

'...I have a reasonable understanding of these matters...'

Strictly speaking, constituent negation, as defined by Klima (1964), should be distinguished from sentential negation with narrow focus on a particular constituent (cf. Willis et al. 2013b: 5–6). However, the actual constructions used to express constituent negation and sentential negation with narrow focus are very frequently identical in the world's languages, and Maltese is no exception, using *mhux* also for the latter, as illustrated in (25).

(25) Korpus Malti v3.0 literature82

Kienu	*qis-hom*	*għaddew*	*xahrejn*	***mhux***	*jumejn.*
AUX.PST.3PL	like-3PL	pass.PFV.3PL	month.DU	NEG	day.DU

'It was as if two months had past, not two days.'

In (24) there is a single, affirmative proposition expressed (that the speaker has a certain kind of idea). In (25) we have, in effect, two conjoined propositions, the first affirmative, the second negative, with the material in the second that is identical to the first unexpressed: it was as if two months had past, (and it

was) not (as if) two days (had passed). From this point of view, the negation in (25) should be seen as sentential negation with ellipsis, not constituent negation, but the close resemblance to actual constituent negation, as in (24), means that the widespread labeling of examples similar to (25) as constituent negation is probably harmless.

5. Negation and indefinite pronouns

The area of grammar discussed in this section is covered in detail by Haspelmath & Caruana (1996), Lucas (2014), and Camilleri & Sadler (2017). Here I just present a brief overview of this rather complex domain. The topic at issue is the interaction of negation with indefinite pronouns; that is, how Maltese expresses meanings such as 'I didn't see anything' or 'No one said anything to anyone'. Like many European languages (including non-standard but not standard varieties of Germanic languages like English), Maltese exhibits a form of what is called, following Labov (1972), negative concord. This is the phenomenon whereby indefinite pronouns in the scope of negation must themselves also be negative (in a sense to be made more precise in a moment). Hence in (26) the presence of the negator *mhux* requires the element translated with English *anything* to be *xejn* and not *xi ħaġa*, which in other contexts would also be translated with English *anything*.

(26) Korpus Malti v3.0 news11479

... l-Partit	*Nazzjonalista*	***mhux***	*qed*	*jeskludi*	***xejn***
DEF-party	nationalist	NEG	PROG	exclude.IPFV.3SGM	nothing

'...the Nationalist Party is not excluding anything.'

In the theoretical literature, items such as *xejn* are either referred to as n-words (following Laka 1990), or, more frequently in recent years, as negative concord items (NCIs). The crucial property of such items is that in contexts such as (26) they appear not to

be negative, since there is no logical double negation with the predicate negator of the kind seen in (27) (= (9) above), whereas in other contexts, such as (28) (= (20) above) and (29), it seems clear that it is *xejn* that is generating the negative interpretation of the clause in which it appears (cf. Giannakidou 2006). Evidence that native speakers of Maltese also consider *xejn* to be inherently negative can be seen from the denominal verb *xejjen* 'to nullify, make nothing'. NCIs thus represent a significant challenge for most compositional theories of natural-language syntax and semantics (see Lucas 2014 for discussion).

(27) Korpus Malti v3.0 news126780 (cf. Stolz 2009: 153)

Ma	*nistgħu-**x***	***ma***	*naħsbu-**x***	*kif...*
NEG	can.IPFV.1PL	NEG	think.IPFV.1PL-NEG	how

'We cannot not consider how...'

(28) Notice observed in Valletta

Tarmi	***xejn***	*hawn.*
throw.IPFV.2SG-NEG	nothing	here

'Don't throw anything here.'

(29) Korpus Malti v3.0 literature20

Tifel	*ta'*	*ħames*	*snin*	*x'-jista'*
child	GEN	five	year.PL	what-can.IPFV.3SGM

jifhem?	***Xejn***.
understand.IPFV.3SGM	nothing

'What can a five-year-old child understand? Nothing.'

As noted above, an indefinite pronoun in the scope of negation that would be translated with English *anything* must be rendered by Maltese *xejn*. This does not mean, however, that *xi ħaġa*, which would typically be translated as *anything* in the question in (30), cannot also appear as the object of a negated verb. But when it does, it is interpreted outside the scope of negation, with a specific indefinite interpretation usually best translated with English *something*, as in (31).

(30) Korpus Malti v3.0 literature58

Taf	*x'-ġara-l-ek?*	*Għand-ek*
know.IPFV.2SG	what-happen.PFV.3SGM-DAT-2SG	POSS-2SG
idea? *Tiftakar*	***xi_ħaġa****?*	
idea remember.IPFV.2SG	something	

'Do you know what happened to you? Do you have an idea? Do you remember anything?'

(31) Korpus Malti v3.0 news137970

Il-Gvern [...],	*jekk*	***mhux***	*se*	*jagħmel*	***xi_ħaġa***
DEF-government	COMP	NEG	FUT	do.IPFV.3SGM	something

urġenti…
urgent

'The government [...], if it doesn't do something urgently, ...'

There are thus two series of indefinite pronouns in Maltese: NCIs, like *xejn*, that are restricted to the scope of negation, and items that appear in other contexts, all of which feature the indefinite determiner *xi*, as shown in Table 2 (taken from Haspelmath & Caruana 1996). Compare the three series of English: *some-*, *any-*, and *no-*, as in *somewhere*, *anywhere*, and *nowhere*.

Meaning	**NCIs**	***xi* series**
Determiner	(l-)ebda	xi
Person	ħadd	xi ħadd
Thing	xejn	xi ħaġa
Time	qatt	xi darba
Place	mkien	xi mkien

Table 2: Maltese indefinite pronouns

Note that in fact the distribution of these items is not quite as neat as Table 2 implies. While only the items in the NCI column have the ambiguous behaviour described above (they look generally negative, except in clauses that contain another expression of negation), Camilleri & Sadler (2017) point out that most of them can nevertheless also occur in certain non-veridical contexts such as questions and conditional clauses with non-

negative meaning, as illustrated in (32), in which we see that *xejn* can also function as a determiner 'any; many'.

(32) Korpus Malti v3.0 news83159

...xtrajt	***xejn***	*ħwejjeġ?*
buy.PFV.2SG	nothing	clothes

'... did you buy many clothes?'

Finally, recall from §3.3.1 that NCIs generally require the *-x* of the bipartite *ma…-x* negative construction to be absent, as illustrated in (33) (= (17)).

(33) Korpus Malti v3.0 academic12

Il-ħajja	*f'-dawn*	*l-istituti*		***qatt***	***ma***
DEF-life	in-DEM.PL	DEF-institution.PL		never	NEG
kienet	*waħda*	*ta'*	*lussu.*		
COP.PST.3SGF	one.F	POSS	luxury		

'The life in those institutions was never one of luxury.'

There are two further details worth noting here. First, the general applicability of this rule of *x*-dropping means that it is surprising that we find *mhux* co-occurring with *xejn* in (26) above. Indeed, it seems that many speakers consider such structures to be ungrammatical (cf. Camilleri & Sadler 2017: 151); but they are robustly attested, albeit as a minority option relative to similar structures with *mhu* (i.e. the expected form with *x*-dropping).[9] There appears to be no similar possibility of *-x* appearing in structures like that in (33), in which a verb (rather than the pronominal copula) is negated with *ma* (not *mhux*) and co-occurs with an NCI. The discrepancy is presumably explained by the fact that, for at least some speakers, *mhux* is felt to be monomorphemic, so that it is either not possible or not necessary to drop the final consonant when it co-occurs with an NCI.

9 A search of Korpus Malti v3.0 with the query "mhux (_PROG|_FUT) _VERB xejn" returns 100 matches, whereas "(m'hu|mhu) (_PROG|_FUT) _VERB xejn" returns 3,199 matches.

The second point is that it is not yet fully clear how closely associated a negative verb and an NCI need to be for the latter to trigger *x*-dropping (cf. Čéplö & Lucas 2020). We can say, informally, that it appears that the two need to be in the same sentence, but this then requires a precise definition of the concept sentence, which will not be attempted here. Note, however, that the NCI and the verb need not be in the same immediate clause: an NCI in a subordinate clause regularly triggers *x*-dropping in a higher clause, as illustrated in (34).

(34) Korpus Malti v3.0 news145572

*...**m'**-għand-ek*	*aptit*	*tagħmel*	***xejn***...
NEG-POSS-2SG	appetite	do.IPFV.2SG	nothing

'You don't feel like doing anything.'

6. Non-negative *-x*

A final observation concerns instances of suffixed *-x* that are sometimes referred to as negative (e.g. Borg & Azzopardi-Alexander 1997: 4), but which are not best analysed as such, either synchronically or diachronically. Recall from §3.1 that the negative suffix *-x*, as well as the indefinite items *xejn* and *xi* (and interrogative *x'*, among other items), are originally derived from the Arabic noun *šayʔ* 'thing'. The process by which this noun came to grammaticalize as a negator is analysed by Lucas & Lash (2010), Lucas (2013; 2018), and Diem (2014), among others. What is relevant to the present discussion is that Arabic *šayʔ* has evolved into numerous different grammatical items in Maltese and the Arabic dialects (cf. Souag 2018). A greater or lesser quantity of the original phonetic material is preserved in the different evolutions (*xejn*, *xi*, *-x/x'*),[10] but there is no necessary link between negation and reduction to /ʃ/. Rather we should envisage one

10 For a discussion of the etymologically unexpected final /n/ in *xejn* and a number of other Maltese items, see Lucas & Spagnol (2022).

major grammaticalization pathway as having been *šayʔ ‘thing’ > *ši (*xi*) ‘at all’. This adverbial element *ši would have been, like the English translation equivalent *at all*, a so-called negative polarity item – that is, restricted in its occurrence to nonveridical contexts such as interrogative, conditional, and negative clauses. In many (but not all) Arabic varieties, including the immediate ancestor of Maltese, this *ši (or a reduced form /ʃ/) was reanalysed in negative clauses as (part of) the expression of negation itself, coming in Maltese to suffice as the sole expression of negation in negative imperative sentences, as discussed in §3.3.1. But the non-negative uses in other nonveridical contexts persisted, sometimes also being reduced just to /ʃ/. The instances of *-x* illustrated in (35) below are thus reflexes of this adverbial *ši ‘at all’ element; they are not synchronically negative (these are not negative clauses), and they do not represent diachronic extensions to non-negative contexts of the negator *-x*. Note that an idiosyncracy of this Maltese reflex of *ši is that its contexts of use have contracted almost exclusively to indirect questions optionally introduced by *jekk* ‘whether’. In such clauses non-negative *-x* has become obligatory, and where these clauses contain *qatt* ‘ever’, as in (35b), *-x* is obligatorily suffixed to *qatt*. Otherwise, it is a suffix on the verb, as in (35a). Non-negative *-x* may, however, also occur in direct questions, and rarely also in conditional clauses, as illustrated in (36) (cf. Wilmsen 2016; Lucas 2018).

(35)

a. Korpus Malti v3.0 news471

Ħafna	*qed*	*jistaqsu*	*jekk*	*wasal-**x***	*iż-żmien*
much	PROG	ask.IPFV.3PL	COMP	arrive.PFV.3SGM-ŠI	DEF-time
li	*ż-żona*	*Schengen*	*tiġi*	*mwaqqfa*	
COMP	DEF-zone	PN	come.IPFV.3SGF	stop.PTCP.PASS.F	

temporanjament.
temporarily

‘Many are asking whether the time has come for the Schengen zone to be temporarily halted.’

b. Korpus Malti v3.0 literature24
staqsie-ni *jekk* *qatt-**x*** *noħlom* *li*
ask.PFV.3SGM-1SG COMP ever-ŠI dream.IPFV.1SG COMP
niżżewweġ.
marry.IPFV.1SG
'He asked me whether I ever dream of getting married.'

(36)

a. Korpus Malti v3.0 literature73
*Intom-**x*** *taraw* *dak* *il-bieb* *f-ix-xellug* *ta-n-niċċa?*
2PL-ŠI see.IPFV.2PL DEM DEF-door PREP-DEF-left GEN-DEF-niche
'Do you see that door on the left of the niche?'

b. Korpus Malti v3.0 news80227
Jekk *qatt-**x*** *ridna* *nkunu* *nafu*
COMP ever-ŠI want.PFV.1PL AUX.IPFV.1PL know.IPFV.1PL
*ngħixu-**x*** *f'-pajjiż* *ta-d-daħq* *jew* *ta-l-biki...*
live.IPFV.1PL-ŠI in-country GEN-DEF-laughter or GEN-DEF-crying
'If we ever wanted to know whether we live in a land of laughter or of tears...'

7. Conclusion

As this article has shown, the major features of the morphosyntax of negation in contemporary Standard Maltese are now relatively well understood. We have seen how negation is expressed anaphorically, at the level of clauses and sub-clausal constituents, and in its interaction with indefinite pronouns. Much work remains to be done, however, before we have a full picture of some of the finer details, including the precise interaction of negation with auxiliaries in verbal chains, and with the different kinds of indefinite pronouns. Another area of particular interest to investigate in future is the extent of dialectal and sociolinguistic variation in the expression of the different kinds of negation in Maltese.

Abbreviations

*	reconstructed form	IPFV	imperfective
1, 2, 3	first, second, third person	M	masculine
ACT	active	MOD	modal
AUX	auxiliary	NEG	negative
COMP	complementizer	OBJ	object (marker)
CONJ	conjunction	PASS	passive
COP	copula	PFV	perfective
DAT	dative	PL	plural
DEF	definite article	PN	proper name
DEM	demonstrative	POSS	possessive
DU	dual	PREP	preposition
EXS	existential	PROG	progressive
F	feminine	PST	past
FOC	focus	PTCP	participle
FUT	future	PURP	purposive
GEN	genitive	SG	singular
IMP	imperative		

Electronic resources

Korpus Malti v3.0: https://mlrs.research.um.edu.mt/

BCv3 (bulbulistan corpus malti v3): www.bulbul.sk/bonito2 (login: guest, password: Ghilm3)

References

Al-Sayyed, Amany & Wilmsen, David (2017) Verbal negation with muš in Maltese and Eastern Mediterranean Arabics. In: Saade, Benjamin & Tosco, Mauro (eds.), *Advances in Maltese linguistics*. Berlin: De Gruyter Mouton, 151–172.

Auwera, Johan van der & Krasnoukhova, Olga (2020) The typology of negation. In: Déprez, Viviane & Espinal, Teresa M. (eds.), *The Oxford handbook of negation*. Oxford: Oxford University Press, 91–116.

Borg, Albert & Azzopardi-Alexander, Marie (1997) *Maltese. Lingua descriptive grammars*. London & New York: Routledge.

Breitbarth, Anne, Lucas, Christopher & Willis, David (2020) *The history of negation in the languages of Europe and the Mediterranean, vol. II: Patterns and processes.* Oxford: Oxford University Press.

Camilleri, Maris & Sadler, Louisa (2017) Negative sensitive indefinites in Maltese. In: Butt, Miriam & King, Tracy Holloway (eds.), *Proceedings of the LFG'17 conferenc*e. Stanford, CA: CSLI Publications, 146–166. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG-2017/

Camilleri, Maris & Sadler, Louisa (2019) The grammaticalisation of a copula in vernacular Arabic. *Glossa* 4(1: 137), 1–33. https://doi.org/10.5334/gjgl.915

Čéplö, Slavomír (2018) *Constituent order in Maltese: A quantitative analysis*. Prague: Charles University Doctoral dissertation.

Čéplö, Slavomír & Lucas, Christopher (2020) *Lanqas*, negative concord and predicate negation in Maltese. In: Čéplö, Slavomír & Drobný, Jaroslav (eds.), *Maltese linguistics on the Danube*. Berlin: Mouton de Gruyter, 177–198 .

Comrie, Bernard (1982) Syntactic–morphological discrepancies in Maltese sentence structure. In: Tasmowski, Liliane & Willems, Dominique (eds.), *Problems in syntax*. Boston: Springer, 3–28.

Dahl, Östen (1979) Typology of sentence negation. *Linguistics* 17, 79–106.

Diem, Werner (2014) Negation in Arabic: *A study in linguistic history*. Wiesbaden: Harrassowitz.

Fabri, Ray & Borg, Albert (2017) Modifiers and complements within the Maltese verb sequence. In: Saade, Benjamin & Tosco, Mauro (eds.), *Advances in Maltese linguistics*. Berlin: De Gruyter Mouton, 67–86.

Giannakidou, Anastasia (2006) N-words and negative concord. In: Everaert, Martin & Van Riemsdijk, Henk (eds.), *The Blackwell companion to syntax, vol. 3*. Oxford: Blackwell, 327–391.

Hansen, Maj-Britt Mosegaard (2013) Negation in the history of French. In: Willis, David, Lucas, Christopher & Breitbarth, Anne (eds.), *The history of negation in the languages of Europe and the Mediterranean, vol. 1: Case studies*. Oxford: Oxford University Press, 51–76.

Haspelmath, Martin & Caruana, Josephine (1996) Indefinite pronouns in Maltese. *Rivista di Linguistica 8*, 213–227.

Horn, Laurence R. (1989) *A Natural history of negation.* Chicago: University of Chicago Press.

Horn, Laurence R. (2001) Flaubert triggers, squatative negation and other quirks of grammar. In: Hoeksema, Jack, Rullmann, Hotze, Sanchez-Valeria, Victor & Van der Wouden, Ton (eds.), *Perspectives on negation and polarity items.* Amsterdam: Benjamins, 173–200.

Jespersen, Otto (1917) *Negation in English and other languages*. Copenhagen: A. F. Høst.

Jespersen, Otto (1933) *Essentials of English grammar*. London: George Allen & Unwin.

Klima, Edward S. (1964) Negation in English. In: Fodor, Jerry A. & Katz, Jerrold J. (eds.), *The structure of language*. Englewood Cliffs, NJ: Prentice-Hall, 246–323.

Labov, William (1972) negative attraction and negative concord in English grammar. *Language* 48, 773–818.

Laka, Itziar (1990) *Negation in syntax: On the nature of functional categories and projections*. Cambridge, Mass.: Massachusetts Institute of Technology PhD dissertation.

Lucas, Christopher (2013) Negation in the history of Arabic and Afro-Asiatic. In: Breitbarth, Anne, Lucas, Christopher & Willis, David (eds.), *The development of negation in the languages of Europe and the Mediterranean, vol. 1: Case studies.* Oxford: Oxford University Press, 399–452.

Lucas, Christopher (2014) Indefinites and negative concord in Maltese: Towards a dynamic account. In: Borg, Albert, Caruana, Sandro & Vella, Alexandra (eds.), *Perspectives on Maltese linguistics*. Berlin: Akademie Verlag, 225–248.

Lucas, Christopher (2018) On Wilmsen on the development of postverbal negation in dialectal Arabic. *Zeitschrift für Arabische Linguistik* 67, 45–71.

Lucas, Christopher & Lash, Elliott (2010) Contact as catalyst: The case for Coptic influence in the development of Arabic negation. *Journal of Linguistics* 46, 379–413. https://doi.org/10.1017/S0022226709990235

Lucas, Christopher & Alluhaybi, Mohammed (2022) The typology of negation across varieties of Arabic. *STUF - Language Typology and Universals* 75(4), 613–641.

Lucas, Chris & Spagnol, Michael (2022) Word-final /n/ in Maltese. In: Turek, Przemyslaw & Julia Nintemann (eds.), Maltese: Contemporary Changes and Historical Innovation. Berlin: Mouton de Gruyter, 47–82.

Meillet, Antoine (1912) L'évolution des formes grammaticales. *Scientia* 12, 384–400.

Payne, John (1985) Negation. In: Shopen, Timothy (ed.), *Language typology and syntactic description: Clause structure*. Cambridge: Cambridge University Press, 197–242.

Peterson, John (2009) "Pseudo-verbs": An analysis of non-verbal (co-)predication in Maltese. In: Comrie, Bernard, Fabri, Ray, Hume, Elizabeth, Mifsud, Manwel, Stolz, Thomas & Vanhove, Martine (eds.), *Introducing Maltese linguistics*. Amsterdam: John Benjamins, 181–204.

Schwenter, Scott (2005) The pragmatics of negation in Brazilian Portuguese. *Lingua* 115, 1427–56.

Souag, Lameen (2018) Arabic-Berber-Songhay contact and the grammaticalization of "thing." In Manfredi, Stefano & Tosco, Mauro (eds.), *Arabic in contact*. Amsterdam: John Benjamins, 53–71.

Spagnol, Michael (2009) Lexical and grammatical aspect in Maltese. In: Stolz, Thomas (ed.), *Ilsienna*. Bochum: Universitätsverlag Brockmeyer, 58–61.

Stassen, Leon (1996) The switcher's paradise: Nonverbal predication in Maltese. *Rivista di Linguistica* 8(1), 275–300.

Stolz, Thomas (2009) Splitting the verb chain in modern literary Maltese. In: Comrie, Bernard, Fabri, Ray, Hume, Elizabeth, Mifsud, Manwel, Stolz, Thomas & Vanhove, Martine (eds.), *Introducing Maltese linguistics*. Amsterdam: John Benjamins, 133–179.

Vanhove, Martine (1993) *La langue maltaise: Etudes syntaxiques d'un dialecte arabe "périphérique."* Wiesbaden: Harrassowitz.

Willis, David, Lucas, Christopher & Breitbarth, Anne (eds.) (2013a) *The history of negation in the languages of Europe and the Mediterranean, vol. I: Case studies.* Oxford: Oxford University Press.

Willis, David, Lucas, Christopher & Breitbarth, Anne (2013b) Comparing diachronies of negation. In: Willis, David, Lucas, Christopher & Breitbarth, Anne (eds.), *The history of negation in the languages of Europe and the Mediterranean, vol. 1: Case studies*. Oxford: Oxford University Press, 1–50.

Wilmsen, David (2016) Polar interrogative -*š* in Maltese: Developments and antecedents. In: Puech, Gilbert & Saade, Benjamin (eds.), *Shifts and patterns in Maltese*. Berlin: De Gruyter Mouton, 175–198.

VARIABLE OVERT MARKING OF PLACE/GOAL WITH PLACE NAMES AS COMPLEMENTS

On the competition between *fi*, *ġewwa*, and *ġo*

Thomas Stolz
& Nataliya Levkovych
& Maike Vorholt

Abstract

The focus of this paper is on the interaction between the spatial prepositions *fi*, *ġewwa*, *ġo* and place names as their complements. On the basis of data extracted from the *Korpus Malti 3.0*, the token and type frequencies of the PPs are presented in order to determine the hierarchical order of the prepositions under scrutiny. Several criteria are checked as to their role in the choice of preposition for the function of expressing Place and Goal. The discussion of the facts is complemented by a preliminary comparison with constructions involving bare place names.

Dan l-istudju jiddiskuti l-interazzjoni bejn il-prepożizzjonijiet spazjali *fi*, *ġewwa*, *ġo* u t-toponimi bħala l-kumplament tagħhom. Permezz ta' data mill-*Korpus Malti 3.0*, tinħareġ il-frekwenza tat-tokens u tat-types tal-Frażijiet Prepożizzjonali biex tiġi determinata l-ordni ġerarkika tal-prepożizzjonijiet analizzati. Numru ta' kriterji jiġu diskussi fid-dawl tal-irwol tagħhom fl-għażla tal-prepożizzjoni biex jesprimu l-Post u l-Iskop. Fid-diskussjoni tar-riżultati isir paragun preliminari ma' binjiet li jinvolvu toponimi waħedhom.

1. Introduction

In his PhD-thesis, Saari (2003) provides substantial proof of the rich and linguistically intriguing phenomenology of Maltese prepositions. Following Saari's lead, Stolz/Ahrens (2017), Schmidt/Vorholt/Witt (2020), and Stolz/Levkovych (2020) explore a variety of aspects of Maltese prepositions which had previously not gained sufficient attention. In these contributions, variation is a recurrent theme which deserves to be further elaborated upon if one intends to describe the grammar of Maltese prepositions comprehensively.

We take up the issue of variation by way of investigating certain patterns in the use of spatial prepositions for which there is ample empirical evidence in contemporary Maltese. We approach the data from the perspective of *Special Toponymic Grammar (STG)* – a concept developed by Stolz/Levkovych/Urdze (2017a–b, 2018) and Stolz/Levkovych (2019a), according to which place names tend to behave differently from common nouns morphosyntactically across languages.[1] In the context of this study, this means that we focus on PPs which express either Place (= *in*, *at*, *on*) or Goal (= *(in)to*) and whose complement is a place name. The PPs we are interested in have the following shape: $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$. What makes this construction type especially intriguing is the multitude of choices speakers of Maltese have when it comes to filling the slot on the left. If Place is the spatial relation to be expressed, it does not seem to make any difference semantically which of the four options of realizing the above construction is chosen. The four options are:

1 The inspiration for STG stems from Nübling/Fahlbusch/Heuser (2015: 64), who propagate the idea that there is a *Special Onymic Grammar* which assumes that proper names generally display properties which are not shared by common nouns.

(a) absence of material exponence (zero-marking) (Stolz/Levkovych 2019b) = bare place name as in (1),
(b) the use of *fi* 'in, at, inside' (Aquilina 1987: 333) as in (2),
(c) the use of *ġew(wa)*[2] 'inside, in, within' (Aquilina 1987: 393–394) as in (3), and
(d) the use of *ġo* 'in, within, inside' (Aquilina 1987: 409) as in (4).

Examples (1)–(4) are drawn from one and the same monograph – an account of the history of Maltese emigration in 19th–20th centuries entitled *L-emigrazzjoni Maltija: Is-seklu dsatax u għoxrin* published in 1999, i.e. an individual author (Edward Attard) gives evidence for the employment of the entire set of options (a)–(d).

(1) Bare place name (Attard 1999: 67)[3]

Dun Guliermu	*qagħad*	***[Ø***	***Sydney]***	*fejn*	*ħadem*	*ħafna*
Dun Guliermu	stay.PFV	**[Ø**	**Sydney]**	where	work.PFV	much

ma-l-Malti-n.
with-DEF-Maltese-PL
'Dun Guliermu stayed **[in Sydney]** where he worked a lot with the Maltese.'

(2) *fi* (Attard 1999: 128)

[F'	***Sydney]***	*Mons. Gonzi*	*kien*	*il-mistieden*
[in	**Sydney]**	Mons. Gonzi	be.PFV	DEF-invite.PPTCPL

ta-l-Kardinal Gilroy…
of-DEF-Cardinal Gilroy
'**[In Sydney]** Mons. Gonzi was the guest of Cardinal Gilroy…'

2 In this study, we gloss over the distinction of long (= *ġewwa*) and short forms (= *ġew*) of this preposition. According to our anonymous reviewers, the latter is only in use with the deictics *hawn* 'here' and *hemm* 'there' yielding *hawn ġew* and *hemm ġew*.

3 In the sentential examples, square brackets mark the boundaries of the PPs under scrutiny. The PPs are additionally highlighted in boldface in the original, the morpheme glosses, and the corresponding translation. Unless otherwise stated, all English translations are ours. In the absence of spatial prepositions the symbol Ø is used.

(3) *ġewwa* (Attard 1999: 127)

waqt	*quddiesa*	*f-il-katidral*	*ta'*	*St Mary's*
during	Holy Mass	in-DEF-cathedral	of	St Mary's
[ġewwa	***Sydney]***...			
[inside	**Sydney]**			

'…during Holy Mass in St Mary's Cathedral **[in Sydney]**…'

(4) *ġo* (Attard 1999: 76)

Skont	*Parnis*	*f-l-1929*	***[ġo***	***Sydney]***	*kien*	*hemm*
according_to	Parnis	in-DEF-1929	**[in**	**Sydney]**	be.PFV	there
madwar		*erba' mija u ħamsin*			*Malti*	
around		450			Maltese	
j-aħdm-u	*f-l-industrija*	*ta-t-tiġieġ*	*u*	*l-bajd.*		
3-work.IPFV-PL	in-DEF-industry	of-DEF-chicken	and	DEF-egg		

'According to Parnis, in 1929 **[in Sydney]**, there were about 450 Maltese working in the chicken and egg industry.'

The place name *Sydney* remains the same for all four examples. The same holds for the spatial relation which is always that of Place. These invariable factors notwithstanding, the construction is realized in four different ways syntactically. In (1), the place name functions as locative complement of the static verb *qagħad* 'stay'. In contrast, examples (2)–(4) involve instances of $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$, which can be classified as spatial adjuncts or adverbials. The PP *ġewwa Sydney* forms part of an NP in (3). Moreover, the PPs occupy different positions within the sentence. In (2), for instance, *f'Sydney* is found sentence-initially whereas the corresponding PPs in (3)–(4) are preceded by another PP (*fil-Katidral ta' St. Mary's* in (3) and *fl-1929* in (4)). The question arises whether syntactic aspects of this kind play a role in the competition between options (a)–(d). This question has yet to be answered. This study is meant to pave the way towards answering the question.

Borg/Azzopardi-Alexander (1997: 157–158) treat the prepositions *fi, ġo,* and *ġewwa* as functional equivalents of each other as markers of Place (interior, at rest) and Goal (interior, motion to). The *Ø-fi=* alternation is the topic of a dedicated case

study by Stolz/Lestrade/Stolz (2014: 225–273), who show that Ø fulfills the same tasks as *fi*. Since *fi* and Ø at the same time compete with *ġewwa* and *ġo* in the same functional domain, there is a complex network of relations which needs to be inquired into in order to determine whether we are dealing with free variation or rule-governed (complementary or partly overlapping) distribution. The guiding question for an investigation that is supposed to study the problem in-depth is whether the options (a)–(d) can replace each other unconditionally in each and every context provided the Ground-NP is a place name.

On account of the above and numerous similar instances of variation, one may ask:

i. whether constructions like those in (1)–(4) are fully synonymous in the first place, and
ii. what the syntactic or other factors are which determine the choice of construction.

Before we can address these questions, it is necessary, however, to understand to what extent the pattern of variation is common in (written) Maltese synchronically. This is exactly what our study is supposed to achieve. For obvious reasons, this paper only marks the beginning of a series of investigations which will explore the subject matter thoroughly. What we say in this study is of a preliminary nature and touches only upon a small selection of interesting phenomena.

To keep our study within reasonable bounds, we do not recapitulate what has been said already as to zero-marking in the grammar of space of Maltese (Stolz/Lestrade/Stolz 2014; Stolz/Levkovych/Urdze 2017b, 2018) and sundry languages (Stolz/Levkovych 2019a). The exclusive topic of the subsequent sections is the synchronically attested alternation of *fi*, *ġewwa*, and *ġo* as material fillers of the prepositional slot in $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$. The range of fillers for the TOP-slot covers almost the entire taxonomy

of place names presented in Nübling/Fahlbusch/Heuser (2015: 206–264) although macronymic settlement names (from villages via cities and countries to continents) are responsible for the bulk of the data. Imagined places such as Harry Potter's *Hogwarts* are also taken on board.

Following this introduction, there are two main sections, viz. Sections 2 and 3. Section 2 is dedicated to the quantitative side of the problem at hand. We have extracted all instances of [PREP$_{\text{PLACE/GOAL}}$ TOP]$_{\text{PP}}$ from the *Korpus Malti 3.0* (https://mlrs.research.um.edu.mt/CQPweb/malti03/) to create a robust data-base for the calculation of type and token frequencies. We show how the different prepositions can be ranked according to their overall frequency and their distribution over different place names (= types). A small selection of qualitative aspects is addressed in Section 3 where we focus on those types which are attested in combination with each of the three contenders. When discussing qualities, we also use sentential examples from the printed version of Attard (1999) – a text that forms part of the input of the *Korpus Malti 3.0*. In Section 4, we draw the conclusions and provide an outlook on potential follow-up studies.

2. Frequencies and shares

In this section, we make use of simple explorative methods to get a first impression of how *fi*, *ġewwa*, and *ġo* relate to each other in terms of quantities. A more sophisticated quantitative account of the data remains a task to be tackled in the future. For practical reasons, we first talk about PREP token frequency in Section 2.1 and then look at TOP type frequency in Section 2.2.

2.1 PREP Tokens

The *Korpus Malti 3.0* yields a turnout of 440,934 tokens for the construction [PREP$_{\text{PLACE/GOAL}}$ TOP]$_{\text{PP}}$. The shares *fi*, *ġewwa*, and *ġo* have

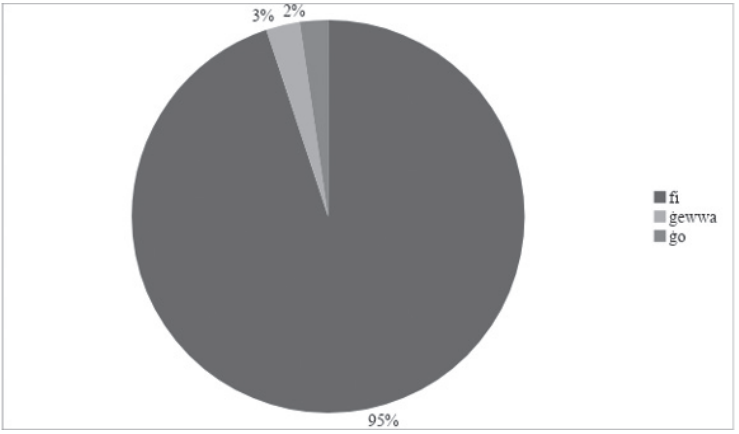


Diagram 1: Shares of *fi*, *ġewwa*, and *ġo* (tokens)

of these types and tokens differ widely. We count 418,450 tokens for *fi* whereas *ġewwa* with 12,274 tokens and *ġo* with 10,210 tokens claim much smaller shares. Diagram 1 reflects these discrepancies by way of showing that *fi* accounts for 95% of all instances of the construction under scrutiny so that the share for *fi*'s competitors *ġewwa* and *ġo* together is as small as 5%.

Given this overwhelming predominance of *fi*, one might want to doubt that there is much of a competition in the first place. Of the three prepositions, *fi* does not only seem to be the default option but also the uncontested majority solution. *Ġewwa* and *ġo* on the other hand, are attested so infrequently that the concept of exception comes to mind. In terms of markedness, *fi* can be classified as the unmarked filler of the PREP-slot in $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$. The other two prepositions, thus, are marked. The frequency of *ġewwa* exceeds that of *ġo* by 2,064 tokens which equals some 20% of the turnout established for *ġo*. We assume that the differences in the token frequencies of the two minor options speak in favour of a markedness hierarchy *fi* > *ġewwa* > *ġo*. In the subsequent Section 2.2, we put this tentative markedness hierarchy to the test in the context of type frequency.

2.2 TOP Types

The TOP type frequencies which result from our search of the *Korpus Malti 3.0* are surprising. There are a total of 816 types, with type being defined as distinct place names. In Table 1, there are four different place names, viz. *Sydney*, *Malta*, *Sqallija* 'Sicily', and *Ruma* 'Rome', which are shown to be compatible with each of the three prepositions *fi*, *ġo*, and *ġewwa*. The PPs are highlighted in boldface. The examples have been extracted from the *Korpus Malti 3.0* on 15 January, 2021.

type	preposition	example
Sydney	*fi*	*għadda tilef ħajtu f'inċident tat-traffiku* ***f'Sydney*** '[] he almost lost his life in a traffic accident **in Sydney** []'
	ġewwa	*Noqgħod* ***ġewwa Sydney*** *f'lokal jismu Merrylands* 'I live **in Sydney** in a place called Merrylands.'
	ġo	*Fil-fatt l-iskema bdiet taħdem* ***ġo Sydney*** *fit-8 ta' Awissu 1967.* 'In fact, the scheme started to work **in Sydney** on 8 August, 1967.'
Malta	*fi*	*minn Settembru 2006 sa Ġunju 2007, introduċejt dawn l-ideat* ***f' Malta*** '[] from September 2006 until June 2007, I introduced these ideas **in Malta** [].'
	ġewwa	*Huwa reat li ġġib annimal* ***ġewwa Malta*** 'It is a crime to bring animals **to Malta** [].'
	ġo	*Jiddikjara li l-każ* ***ġo Malta*** *għalih hu magħluq* 'He declares that **in Malta**, the case is closed for him [].'
Sqallija	*fi*	*b'hekk l-għadu tal-Maltin niżel* ***fi Sqallija*** '[] in this way, the enemy of the Maltese landed **in Sicily**.'
	ġewwa	*B'suċċess ieħor fi triathlon* ***ġewwa Sqallija****, huwa jinsab f'forma eċċezzjonali* 'With another success at the triathlon **in Sicily**, he finds himself in eccellent form [].'
Sqallija/ Ruma	*ġo*	*għandek żewġ organizzazzjonijiet, waħda* ***ġo Sqallija*** *u l-oħra* ***ġo Ruma*** '[] you have two organizations, one **in Sicily** and the other **in Rome** [].'
Ruma	*ġewwa*	*L-idea illi jitwaqqfu dawn is-seminarji bdiet* ***ġewwa Ruma*** 'The idea to establish these seminars began **in Rome** [].'
	fi	*meta l-Faxxisti ħatfu l-poter f'idejhom* ***f'Ruma*** '[] when the Fascists took the power into their hands **in Rome** [].'

Table 1: Four selected place names (types) as complements of *fi*, *ġewwa*, and *ġo*

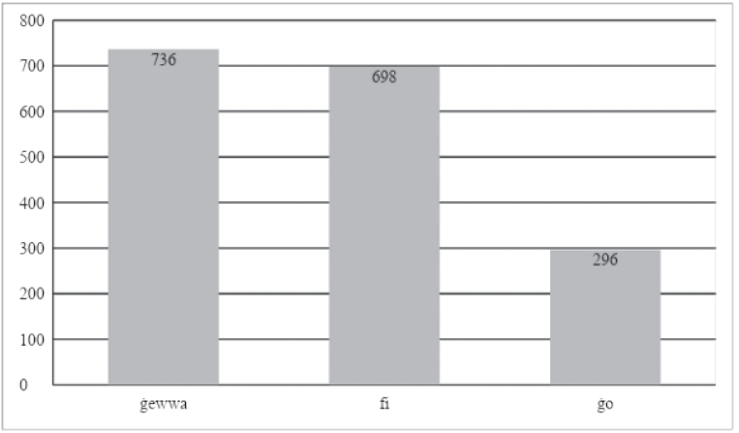


Diagram 2: Type frequencies of *fi*, *ġewwa*, and *ġo*

It strikes the eye that each of the place names combines freely with each of the prepositions. The question arises whether this pattern of unrestricted compatibility has repercussions on the side of the quantities in terms of types per preposition. As shown in Diagram 2, *fi* does not claim the biggest number of types as one might expect on account of *fi*'s token frequency.

Fi is only second best because it is ousted by *ġewwa*. In contrast to *fi* and *ġewwa*, *ġo* has a comparatively low type frequency – a fact that fits in nicely with *ġo*'s subordinate position on the above markedness hierarchy. This hierarchy is violated by *fi* and *ġewwa* because their type frequencies fail to replicate the ranking order established on the basis of the token frequencies.

In point of fact, none of the three prepositions occurs with the entire set of place names. With 736 of 816 types, *ġewwa* is attested in combinations with 90% of all place names. As to *fi*, the 698 types correspond to combinations with 86% of all place names. The combinability of *ġo* and place names is much more limited since the 296 types equal combinations with 36% of all place names. Put differently, *fi* yields an extraordinarily high token frequency although this preposition fails to combine with 14% (= 118) of the place names. We cannot rule out the possibility that this

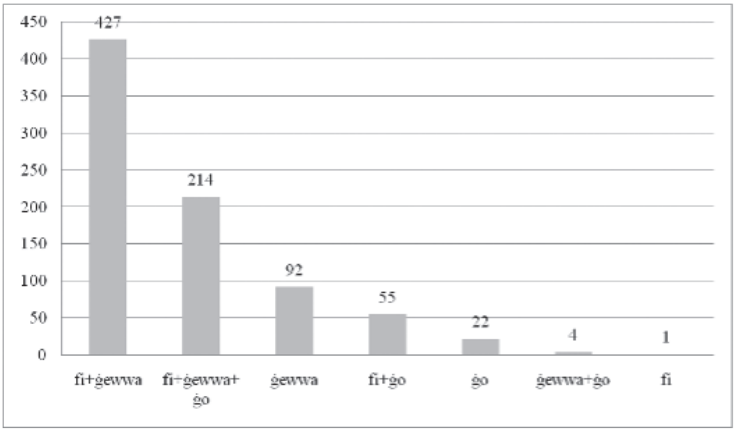


Diagram 3: Type frequency of multiple/single combinability

result is the effect of the composition of the corpus. Before we can draw any conclusions, further research is necessary to determine whether the above frequencies are statistically significant.

Further interesting aspects come to the fore when we look at the possibility of types combining with several of the prepositions. It is clear from Diagram 3 that there is a sizable majority of place names which combine with two or all three of the prepositions like those featured in Table 1. Binary and ternary combinations together cover 700 place names (= 86%). The preferred combination is *fi/ġewwa.* It accounts for 61% of all cases of multiple combinations. What strikes the eye particularly, however, is the minimal number of place names which is attested exclusively in combination with *fi*. There is only a single place name of this kind, i.e. *Mount Carmel* which is attested 157 times as complement of *fi*. In contrast, *ġo* counts twenty-two times as many monopolized place names. *Ġewwa* ranks on top because this preposition counts four times as many monopolized place names as *ġo* and ninety-two times as many as *fi*.

If a given place name is attested as complement of *fi*, the probability is high that the same place name is also attested with

at least one of the other two prepositions. The probability is much lower in the case of *ġo* and even more so in the case of *ġewwa*. The supposed unmarked status of *fi* is thus challenged once again by *ġewwa*. The former is the winner of the competition only in the domain of token frequency whereas the latter is at its strongest in the domain of type frequency.

We complement the above observations by way of discussing the token frequencies of those types which are compatible with several prepositions. We compare the different turnouts of individual PREP-TOP combinations to determine whether they always reflect the same hierarchical order of the prepositions. Six logical possibilities exist, namely (with ≥ meaning 'at least equally frequent as'):

(i) *fi* ≥ *ġewwa* ≥ *ġo*
(ii) *fi* ≥ *ġo* ≥ *ġewwa*
(iii) *ġewwa* ≥ *fi* ≥ *ġo*
(iv) *ġewwa* ≥ *ġo* ≥ *fi*
(v) *ġo* ≥ *fi* ≥ *ġewwa*
(vi) *ġo* ≥ *ġewwa* ≥ *fi*

We take account of all those place names which combine with at least two of the prepositions. The working hypothesis assumes that wherever *fi* competes with another preposition the token frequency of *fi* is not surpassed by that of another preposition. Similarly, *ġewwa* is attested at least as frequently as *ġo*. This means that we expect evidence of (i.) and (iv.) (when *fi* fails to show up) but not of (ii.)–(iii.), (v.)–(vi.). The hypothesis is only partly borne out by the facts.

Diagram 4 shows that the ranking order of type (i.) is indeed the most frequent option. With 581 types it covers 83% of the 698 types of those place names which combine with several prepositions. Two further possible types had to be removed because they are attested in incomplete fragments of sentences only. The ranking

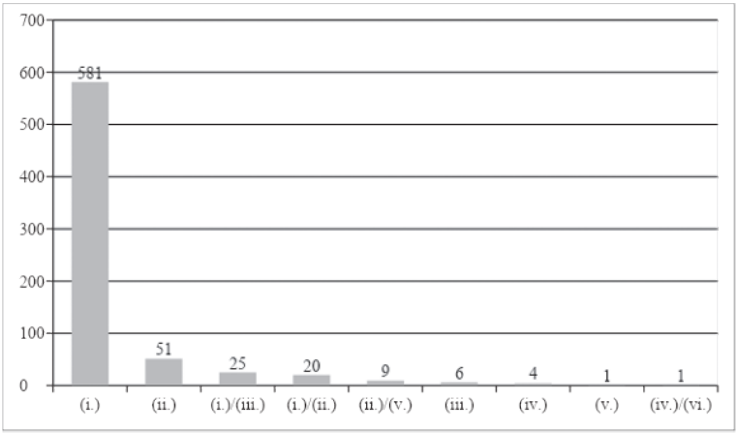


Diagram 4: Type frequency of ranking order

order type (iv.), on the other hand, is only marginally attested with its type frequency of four (= 0.5%). In contrast, the ranking order types (ii.)–(iii.), whose realization has been excluded, show up in our data. It is true that (iii.), has only a share of 0.8%, but (ii.) claims a share of 7% (with 51 types) and thus covers a sizable number of cases. For 55 types (= 8%), it cannot be decided which of two possible ranking orders is the case because the prepositions display identical token frequencies. Almost as predicted the ranking orders (v.)–(vi.) are attested in negligible numbers.

Fi is more frequent than *ġewwa* in combinations with 633 different types whereas the opposite, i.e. *ġewwa* being more frequent than *fi*, is confirmed only for eleven types. It is remarkable that the frequency of *ġo* exceeds that of *ġewwa* with 61 types (it is the inverse for 631 types). *Ġo* outranks *fi* in six cases whereas *fi* is more frequent than *ġo* with 692 types. Therefore, *fi* has a particularly strong position if token frequency is taken account of. Neither *ġewwa* nor *ġo* can compete with the dominant role of *fi*. Nevertheless, the picture is not as straightforward as expected since 9% (= 66) of the place names which combine with several prepositions deviate from the predicted patterns.

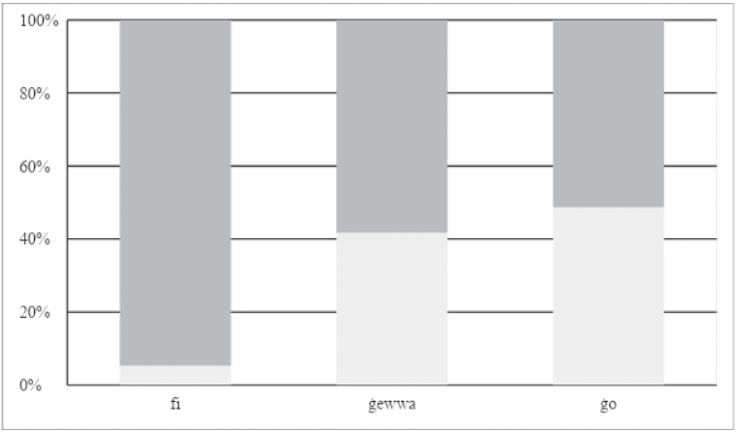


Diagram 5: The role hapaxes per preposition

The average token frequency per type is 540 for the entire data set. *Fi* is the only preposition to exceed this value. The average token frequency per type as complement of *fi* is up to 599. The averages calculated for *ġewwa* and *ġo* are 17 and 34 tokens per type, respectively. What is more, for both *ġewwa* and *ġo* more than 40% of the types display the minimum token frequency of one. In the case of *fi*, only 5% of all types are equally infrequent. These differences can be gathered from Diagram 5. We take the differences to add further support to the hypothesis that *fi* is the default preposition.

It is worth noting that of the 22 types for which *ġo* holds the monopoly, twenty (= 91%) are attested only once and the remaining two types have a token frequency of two. In contrast, the sole example of a place name that exclusively combines with *fi* yields a turnout of 159 tokens. In the case of *ġewwa*, 76 of 92 monopolized types (= 83%) are single occurrences. However, the token frequencies of the remaining sixteen types range from two to fourteen. *Fi* practically fails to monopolize on a grand scale but its sole monopoly is firmly established. *Ġo*, however, boasts a sizable number of monopolies without reaching token frequencies which

prove that these monopolies are watertight. *Ġewwa* represents a kind of in-between case. For the vast majority of its monopolized types a characterization similar to that of *ġo* is fitting. There is, however, a minority of monopolized types of *ġewwa* whose token frequencies are suggestive of a certain degree of robustness of the monopoly.

We close this section with a glance at the types with the highest token frequencies. To this end, we select the ten top ranking types for each of the three prepositions. We present the results for each preposition separately in Table 2–4. The place names are given in normalized spelling, glossing over the, at times, rather numerous alternative spellings (including obvious typos). To facilitate comparison, the token frequency of a given type is additionally computed for the competing prepositions in the two columns on the right. Grey shading identifies those place names which are featured on all three of the top-ten lists. Double underlining marks a type which will be in the focus of the discussion of Section 3.

It comes as no surprise that the two topmost ranks of each preposition host the names of the two major islands of the Maltese archipelago, namely *Malta* and *Għawdex* 'Gozo'. The importance the EU has for the political, economic, and social developments in Malta is reflected by the recurrence of the type *Brussels* in Tables 2–4. Three things need to be mentioned in connection with these types. The place name *Malta* is referentially ambiguous as it can refer either to the island of Malta or to the country – the Republic of Malta. Similarly, Brussels has at least two readings, namely as settlement name referring to the city of Brussels and as synecdoche for the EU and its administrative institutions. Thirdly, it is interesting to see that *Għawdex* outranks *Malta* in Table 3, i.e., *ġewwa* prefers combinations with *Għawdex* over those with *Malta* by a ratio of almost two-to-one. Except *il-Belt* 'Valletta', place names like *l-Ewropa* 'Europe' are presented without definite pro-clitic.

rank	**type**	***fi***	***ġewwa***	***ġo***
1	*Malta*	98,759	756	180
2	*Għawdex*	37,269	1,470	139
3	*Ewropa*	17,608	154	3
4	*Brussels*	14,386	299	38
5	*Libja*	7,435	152	28
6	*Ingilterra*	6,388	284	8
7	*Italja*	6,157	139	4
8	*Valletta*	5,330	2	5
9	*Rabat*	5,268	170	8
10	*Marsa*	4,936	124	7

Table 2: Top ten types for *fi*

rank	**type**	***ġewwa***	***ġo***	***fi***
1	*Għawdex*	1,470	139	37,269
2	*Malta*	756	180	98,759
3	*il-Belt*	541	68	648
4	*Brussels*	299	38	14,386
5	*Ingilterra*	284	8	6,388
6	*Bormla*	250	38	2,893
7	*Paceville*	231	38	1,658
8	*Wied X*	198	76	863
9	*Londra*	186	20	2,680
10	*Rabat*	170	8	5,268

Table 3: Top ten types for *ġewwa*

rank	type	*ġo*	*ġewwa*	*fi*
1	*Malta*	180	756	98,759
2	*Għawdex*	139	1,470	37,269
3	*Wied X*	76	198	863
4	*il-Belt*	68	541	648
5	*Ħal Qormi*	40	142	3,440
6–8	*Bormla*	38	250	2,893
6–8	*Brussels*	38	299	14,386
6–8	*Paceville*	38	231	1,658
9	*Franza*	33	88	3,111
10	*Pembroke*	29	151	1,591

Table 4: Top ten types for ġo

Except *Wied X* in Table 3–4, all types are genuine cases of macro-toponymical settlement names in the broad sense of the term. The *X* in *Wied X* and other place-name types is a variable for obligatory but variable second constituents. We address the problems posed by *Wied X* together with those mentioned above in connection with *Malta*, *Għawdex*, and *Brussels* in Section 3. To wrap up this section, we emphasize that all of the top-ten types are compatible with each of the three prepositions. The ubiquitous preponderance of *fi* notwithstanding, the possibility of *fi*, *ġewwa*, and *ġo* taking identical complements such as *Brussels* renders it unlikely that the choice of preposition depends solely on (yet to determine) inherent properties of the place name.

3. Motivating frequencies

In this section we pose the question whether the above frequencies are motivated by factors yet to be discovered or

the incidental result of free variation. For reasons of space, we restrict the discussion to four criteria, all of which can be considered properties (of different kinds) of the place names involved in $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$. Some properties of the prepositions are discussed, too, along the way. In Section 3.1, we check to what extent the parameter of familiarity (= local vs foreign place names) is relevant for the choice of preposition. Section 3.2 looks into the possibility that the ontological class to which the place name's referent belongs determines which of the prepositions is favoured. The distinction of simple vs complex place names is the topic of Section 3.3. The PP $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$ is reviewed in combination with the motion verb *wasal* 'arrive' on the basis of the data found in Attard (1999) in Section 3.4.

3.1 Familiarity

There are two different ways of defining the familiarity of place names. What comes to mind first is the distinction of place names referring to local geo-objects (= in Malta) vs those which refer to geo-objects abroad (= outside of Malta). The second possibility distinguishes place names coined in Maltese and those which are used in the shape they have in a foreign language. For a start, we look at the distinction of geo-objects in and outside of Malta.

3.1.1 *Places inside and outside of Malta*

Of the 816 types, 635 (= 78%) refer to geo-objects which are situated beyond the boundaries of the Republic of Malta. Only 179 types (= 22%) have a local reference. Two types could not be classified because they are ambiguous as to the places they refer to (in Malta or somewhere else). Accordingly, the three prepositions boast a majority of types with reference to geo-objects outside of Malta. The majority ranges from 65% for *ġo* via 77% for *fi* to 78% for *ġewwa*, as shown in Diagram 6. The results for *fi* and

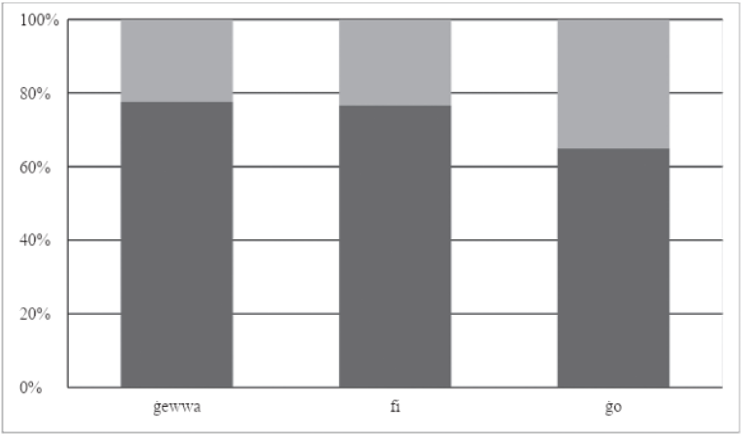


Diagram 6: Local vs foreign geo-reference of types per preposition

ġewwa correspond more or less to the overall picture. This is different with *ġo*, whose shares for foreign and local place names do not match the above percentages since local place names are overrepresented in comparison to their shares in the domain of the two other prepositions. It is also worth mentioning that *ġo* is attested in combinations with 58% of the 179 types with local geo-reference although *ġo*'s share of the entire set of 816 types is only 36%. (Both *fi* and *ġewwa* boast shares which exceed 90% of the 179 types with local reference.) On this basis, one might assume that familiarity is a factor when *ġo* is chosen as filler of the PREP-slot in the construction type under investigation.

To test the tenability of this assumption, we first have a look at the twenty-two types for which *ġo* has the monopoly. The types are presented in Table 5. Except for *Qasam* and *Ramla*, each with a token frequency of two, all entries in Table 5 are attested only once in the *Korpus Malti 3.0*. For the foreign place names, the country which hosts the geo-object that is referred to is identified in brackets.

category	type	sum
foreign	*Cesenatico* (Italy), *Coburg* (Germany), *Dafur* (Sudan), *gżira X* (Spain), *Kisubi* (Uganda), *Lisboa* (Portugal), *Livigno* (Italy), *Maroubra* (Australia), *Mechelen* (Belgium), *Patras* (Greece), *Potsdam* (Germany), *Ravello* (Italy), *Rosemead* (USA), *Saint Pierre et Miquelon* (France), *Susa* (Iran), *Tamworth* (UK), *Transnistria* (Moldavia), *Wuppertal* (Germany)	18
local	*Qammieh, Qasam, Ramla, Ramla ta' Ċirkewwa*	4
total		22

Table 5: Types monopolized by *ġo*

Table 5 does not support the hypothesis according to which familiarity is important for the choice of *ġo* since 82% of the types mentioned in Table 5 refer to geo-objects outside of Malta. The picture changes visibly if we turn our attention to the types with particularly high token frequencies ($n \geq 10$) for combinations with *ġo*. Among the top-ten types ranked in Table 4 we know that there are only two place names – *Brussels* and *Franza* – which have a geo-reference outside of Malta. All other items in Table 4 are local place names. Tables 2–3 show that types with foreign geo-reference are more common on the top-ten ranks of *fi* and *ġewwa*. In the former case, there are five types of this kind (*Brussels*, *Ewropa*, *Ingilterra*, *Italja*, *Libja*), i.e., half of the top-ranking types refer to geo-objects beyond the borders of Malta. As to *ġewwa*, we notice three types with foreign reference – *Brussels*, *Ingilterra*, *Londra*. *Ġo* on the other hand yields the smallest number of types of this kind in Table 4. As results from the types in Table 6, local place names constitute the majority also between ranks 11–34. Types which refer to places outside Malta are highlighted in boldface.

rank	type	tokens
11	*Ħaż-Żabbar*	22
12–13	***Londra***, *Żebbuġ*	20
14–15	*Ħal-Luqa, Tas-Sliema*	19
16	*Marsaxlokk*	18
17	*Kalkara*	17
18	*Ħal Farruġ*	16
19–20	*Birkirkara,* ***New York***	14
21–22	***Pariġi***, ***Sqallija***	13
23–25	*Birgu, Ħal Far, San Ġwann*	12
26–30	*Birżebbuġa, Furjana, Gżira Għawdxija, Kastilja,* ***Spanja***	11
31–34	*Ħal Tarxien,* ***Ruma***, ***Strasburgu***, *Tigne*	10

Table 6: Ranks 11–34 for *ġo*

Local places outnumber foreign places by a ratio of three-to-one. Except *Londra*, no place name with foreign geo-reference exceeds the token frequency of fourteen. If we compare these results with those of Table 5, where foreign place names are clearly dominant quantitatively, we understand that names of local places and *ġo* go together very well. This impression receives further support if we take the corresponding data for the other two prepositions into account.

On ranks 11–34, *ġewwa* provides evidence of seven names of foreign places as opposed to seventeen with local geo-reference whereas *fi* divides the twenty-four rank positions evenly between twelve names of foreign places and twelve which refer to local places. This means that the situation is very similar to that reported for the top ten, namely, half of all place names which combine with *fi* refer to objects outside of Malta. In the case of *ġewwa* and *ġo*, the share of names of foreign places shrinks to about a third of all entries with *ġo* yielding the lowest figures.

The degree of familiarity[4] speakers have with the place to which the complement of the PP refers possibly influences their choice of preposition although only weakly since *ġo* is compatible with names of foreign places and nothing seems to prevent the other two prepositions from taking names of local places as their complements. Thus, we are not talking about strict rules and yes/no-decisions. It is not even entirely clear whether what we have observed here can be termed a preference. We assume that familiarity is only one among a variety of factors which channel the choice of preposition. Possibly they are strongest when they conspire, in a manner of speaking.

3.1.2 *Aliases*

The *Korpus Malti 3.0* provides ample evidence of the co-existence of referentially identical alternative place names. We open the discussion with a particularly striking case for which we can refer back to Table 2–4. It is to be expected that among the top-ten types of the three prepositions the capital city of Malta is represented, too. This is the case indeed but, most interestingly, *fi* clearly favours the internationally known place name *Valletta* whereas the two other prepositions prefer the Maltese coining *il-Belt* as shown in Table 7.

type	***ġo***	***ġewwa***	***fi***	**sum**
Valletta	5	2	5,330	5,337
il-Belt	68	541	648	1,257
totals	73	543	5,978	6,594

Table 7: Synonyms for Malta's capital

4 The reviewers emphasize that the concept of familiarity needs to be elaborated upon further because not every speaker of Maltese might be familiar with every local place name whereas for the same native speakers of Maltese, certain foreign place names might be absolutely familiar. On account of this valuable comment, we envisage to look into this problem more deeply in a follow-up study.

The discrepancy is striking. *Fi* accounts for nearly 100% of all tokens of *Valletta* filling the top-slot in the construction under review. In contrast, the share of *fi* drops to 52% if we replace *Valletta* with *il-Belt*. In the latter case, *ġewwa*'s share is up to 43% and that of *ġo* increases to 5%. On the top-ten lists of *ġewwa* and *ġo*, *il-Belt* occupies ranks 3 and 4, respectively (see Table 3–4) whereas *Valletta* does not even show up on ranks 11–34. Of all *ġo*-PPs which refer to Malta's capital city, 93% involve *il-Belt*. In the case of *ġewwa*, *il-Belt* occurs in 99% of all PPs. As to *fi*, *Valletta* is on rank 8 (see Table 2) whereas *il-Belt* winds up 105 ranks further down, i.e. on rank 113. Only 11% of the *fi*-PPs which refer to the capital of Malta take *il-Belt* as their complement. There is thus a clearly discernible divide between *fi* on the one hand and *ġo* and *ġewwa* on the other.

To check whether the above situation is an idiosyncrasy of the synonymous names of the capital,[5] we look at the traditional name pairs of the so-called *Cottonera* or *Three Cities*, viz. *Bormla* = *Cospicua*, *Senglea* = *Isla*, and *Birgu* = *Vittoriosa*. In Table 8, we confront the token frequencies of the competing place names with each other to see whether the choice of name has an effect on the choice of preposition.

city	type	*ġo*	*ġewwa*	*fi*	sum
A	*Birgu*	12	120	2,000	2,132
	Vittoriosa	0	2	16	18
B	*Bormla*	38	250	2,893	3,181
	Cospicua	0	1	17	18
C	*Isla*	1	64	971	1,036
	Senglea	0	0	0	0
totals		51	437	5,897	6,385

Table 8: Token frequency of alternative place names (Cottonera)

5 Owing to the wide margin of referential ambiguities, we have not been able to run the test for the pair of names which is used for the island capital of Gozo which comes as either *Rabat (t'Għawdex)* or *Victoria*. Since the bracketed addition is often absent from the hits found in the *Korpus Malti 3.0*, it cannot be ruled out that reference is to Rabat in Malta. Similarly, the place name *Victoria* is not distinctive since it frequently refers to the Australian state of Victoria or occasionally to other cities of this name in various countries worldwide.

The grey-shaded cells of Table 8 host zeros. The place name *Senglea* is not attested at all as filler of the TOP-slot in $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$. Moreover, *ġo* does not combine with either *Cospicua* or *Vittoriosa* and is attested only once in combination with *Isla*. In contrast, the *ġo* takes *Birgu* and especially often, also *Bormla* as complements. The data are not absolutely conclusive but one might argue that *ġo* prefers combinations with the older layer of place names over combinations with the relatively recent alternative place names whose Italian origin is still transparent (whereas that of *Birgu* < Italian *borgo* 'village, suburb' (Aquilina 1987: 125) is opaque). Note that this pattern also holds for *ġewwa* and *fi*. What is different with these two prepositions as opposed to *ġo* is the frequency with which they take *Isla* as complement. Since all three of the prepositions behave similarly, the results are different from those mentioned in the contexts of the competition of *il-Belt* and *Valletta*.

Table 9 presents ten pairs of synonymous place names which refer to places outside of Malta. As in the previous table, the token frequencies of the alternatives are directly compared to each other.

place	**type**	***ġo***	***ġewwa***	***fi***	**sum**
A	*Antwerp*	1	2	17	20
	Anversa	0	1	0	1
B	*Corsica*	1	0	12	13
	Korsica	0	1	0	1
C	*Denmark*	1	0	7	8
	Danimarka	0	19	643	662
D	*Iceland*	0	2	4	6
	Islanda (Iżlanda)	0	6	268	274
E	*USA*	0	3	36	39
	Istati Uniti	0	60	4,775	4,835

F	*Sweden*	2	0	33	35
	(L-I)Żvezja	0	28	749	777
G	*Jerusalem*	0	4	1	5
	Ġerusalemm	1	1	191	193
H	*Marseille*	1	1	12	14
	Marsilja	2	2	87	91
I	*Netherlands*	0	1	30	31
	Olanda	0	21	1,054	1,075
J	*St Petersburg*	0	1	7	8
	San Pietruburgu	0	1	32	33
totals		9	154	7,958	8,121

Table 9: Token frequency of alternative place names (international)

There are altogether eighteen zeros, thirteen of which are found in the *ġo*-column. The token frequencies are generally low for *ġo*. Only *Sweden* and *Marsilja* are attested twice each, all other cases are hapaxes. This means that the preposition *ġo* does not contribute substantially to the frequency of any of the alternatives. The ten name-pairs generally display preferences for a relatively robust majority option over an, at times, very infrequent minority option. Except in the case of pairs B, G, and J, *ġewwa* behaves in accordance with *fi* in the sense that the former has the same preferences as the latter. The underrepresentation of *ġo* in Table 9 is no surprise since we are dealing with place names which refer to geo-objects outside of Malta. In corroboration of the tendency observed in connection with the name pairs in Table 8, Table 9 shows a clear preference for those place names whose shape looks Maltese. This is the case in eight out of ten pairs in Table 9 namely in C–J. However, the clear preference for a given type does not imply a similarly clear preference for a given preposition other than the expected dominance of *fi*.

3.2 Ontological class

As mentioned in the introductory section, the three prepositions are considered to be largely synonymous. Aquilina's (1987: 333, 393–394, 409) English translations of *fi*, *ġewwa*, and *ġo* share two English translation equivalents, namely *in* and *inside*. In addition, English *within* is also mentioned as a possible translation of *ġewwa* and *ġo.* For *fi* alone, the dictionary also lists English *at*. What complicates things somewhat is the different order in which the English translation equivalents are presented for the prepositions. In the case of *fi,* the translation candidate *in* comes before *at*, which in turn comes before *inside*, whereas *inside* is a translation equivalent #1 for *ġewwa* followed first by *in* and then by *within*. Finally, *ġo* resembles *fi* insofar as its first English translation is *in* and the third option *inside* with *within* sandwiched between these two. Superficially, this variable order of possible English translations looks random. But might it be suggestive of perhaps only very subtle meaning differences? For the sake of the argument, we proceed on the assumption that these subtle meaning differences are such that they render a given preposition particularly suited for combinations with place names whose geo-referents belong to certain ontological classes.

In Section 2.2, we argued that place names like *Malta* are generally ambiguous because they can refer either to an individual island or to the country the island forms part of. In the same section, we also saw, in connection with the topmost positions of the hierarchies in Tables 2–4, that the island name *Għawdex* surpasses the ambiguous place name *Malta* on the list of the top-ten types of *ġewwa*. On this basis, we compare the token frequencies of those island names which are identical with the name of a state to that of names which exclusively refer to a particular island. This is done summarily in Diagram 7 (excluding *Malta* and *Għawdex*).

There are thirty types, eleven of which are ambiguous. As comes to the fore in Diagram 7, more than half (53%) of all *fi-*

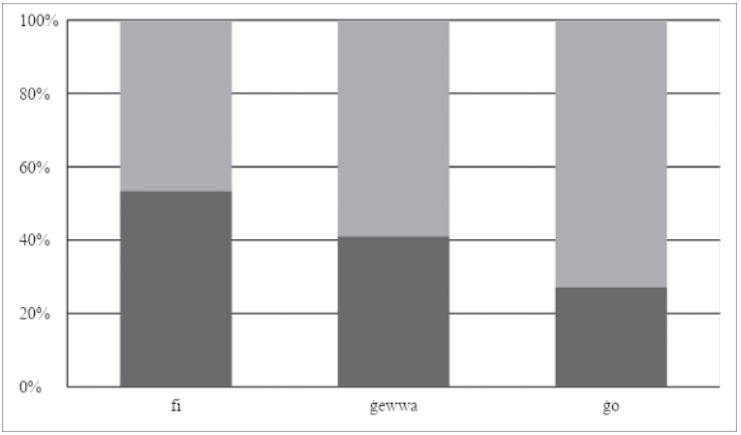


Diagram 7: Ambiguous vs clear reference with island names (tokens)

PP tokens involve ambiguous place names (such as e.g. *Jamaica*, *Tajwan*, and *Ċipru* 'Cyprus'). The share of this class of place names is down 41% with *ġewwa* and 27% with *ġo*. Except in the case of *Saint Pierre et Miquelon*, which is attested only once (in combination with *ġo*), *fi* ousts its contestants quantitatively in each case. Yet, *ġewwa*, and much more so *ġo*, have preferences different from those of *fi*. Island names like *Kemmuna* 'Comino', *Manoel Island*, and *Sqallija* 'Sicily' are responsible for much of the turnout of the two minor prepositions.

Since island names which are identical to state names more often than not refer to places outside of the Maltese archipelago, we are probably facing the effect of two factors, namely, reduced familiarity and potential ambiguity, which conspire to further the cause of *fi*, to put it this way. What we cannot confirm, however, is the systematic use of different prepositions to disambiguate potentially ambiguous island names. A case in point is *Ċipru* 'Cyprus', which is the complement of *ġo* eight times and that of *ġewwa* twenty-one times (as opposed to 931 tokens which go to the credit of combinations with *fi*). Not only is it difficult to decide for the analyst whether the place name has the one or the other reading, but the evidence of *Ċipru* referring to the island and not

(also) to the Republic (or former Crown Colony) of Cyprus is scarce. Of the twenty-four tokens which do not stem from *fi*-PPs, only three invite an interpretation of the island-kind. There is thus no general rule according to which the ambiguous complement of every *ġo*-PP or *ġewwa*-PP must refer to the island and not to the country.

The ontological class of the geo-object does not seem to be a very strong factor when it comes to choosing a preposition. This factor is also associated with some of the cases to be discussed in the next section.

3.3 Structural complexity

For this study, we adopt a simple definition of structural complexity. The criterion is applied exclusively to the filler of TOP-slot of $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$. The place name is termed structurally simple if it consists of a single word. It is considered to be structurally complex if it comprises of at least two words. We are aware of the fact that the three prepositions display different degrees of complexity in terms of segmental length, for instance. *Fi* may be reduced to *f'* if the following word starts with a vowel or a single consonant, *ġewwa* boasts a truncated allomorph *ġew*, which we have not distinguished from its long equivalent in this study. The question to what extent the structural properties of the prepositions themselves are crucial for the phenomenon at hand needs to be answered in follow-up studies.

Does the simple-complex distinction correlate with the choice of preposition? Typical representatives of complex place names are hagionymic place names, i.e. places named after Christian saints, because these place names often have a binary structure with an initial element *San(tu/ta)* which normally cannot be dropped. There are twenty types with this structure, twelve of which refer to places outside of Malta and eight have a local geo-reference. In some cases, the exact geo-reference is doubtful because there are

several places which go by the same name. Table 10 informs us about the relevant token frequencies. The table is divided in two: the upper part hosts the international cases whereas the bottom part is reserved for the place names which refer to places in Malta or Gozo. The place names which are attested in *ġo*-PPs are placed at the opposing ends of the table. Empty cells are shaded grey. The order in which the types are listed is meant to facilitate recognizing the special behavior of *ġo*.

country	**type**	***ġo***	***ġewwa***	***fi***	***sum***
Panama	*San Blas*	2		133	135
Brazil	*Sao Paolo*	1		24	25
France	*Saint Pierre u Miquelon*	1			1
Italy	*San Gregorio da Sassola*		1		1
Italy	*San Benedetto (del Tronto)*		2	2	4
Italy	*San Pietru*		1	2	3
Russia	*San Pietruburgu*		1	32	33
San Marino	*San Marino*		5	53	58
USA	*San Antonio*		1	15	16
USA	*San Diego*		2	24	26
USA	*San Francisco*		4	54	58
USA	*Santa Monika (USA)*		1	3	4
Malta	*San Lawrenz*		8	237	245
Malta	*San Luqa*		17	87	104
Malta	*San Tumas*		1	13	14
Malta	*San Ġiljan*	7	82	2,148	2,237
Malta	*San Ġwann*	12	39	1,141	1,192
Malta	*San Pawl il-Baħar*	6	39	2,553	2,598
Malta	*Santa Luċija (Malta)*	1	36	502	539
Malta	*Santa Venera (Malta)*	6	52	1,253	1,311
totals		36	292	8,276	8,604

Table 10: Hagionymic place-name types in $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$ (tokens)

It is obvious that all three of the prepositions are compatible with this class of place names. On closer inspection however, it comes to the fore that the prepositions behave differently. The higher overall frequency of PPs involving local place names as complements meets our expectations. What is remarkable nevertheless is connected to *ġo*. This preposition is only marginally attested in combinations with place names of the hagionymic kind which refer to geo-objects abroad. Only three of twelve types allow for combinations with *ġo*. As to place names which refer to places in Malta, *ġo* is much more common as head of $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$. Five out of eight types combine with *ġo*. Furthermore, all of the place names with a Maltese geo-reference also occur in PPs with *ġewwa* and *fi*, which is not the case with place names referring to geo-objects elsewhere in the world. On account of this differential behaviour of the prepositions, one might conclude that it is not so much structural complexity but familiarity which is crucial for the use of the prepositions. *Ġo* prefers combinations with place names with local geo-reference over those which refer to places in other countries.

The structural complexity of place names is frequently connected to the presence of a so-called classifier (Anderson 2007: 186). The classifier normally identifies the ontological class to which the named geo-object belongs – or originally belonged because in the course of their existence place names may dissociate from the geo-object they were coined for (Nyström 2016). For brevity's sake, we focus on a typical classifier in the Maltese context, namely the village classifier *Ħal*. Place names which contain this classifier do not normally refer to places outside of Malta so that gradual differences on the parameter of familiarity can be counted out as factors.

Table 11 gives a quantitative account of the combinations of the three prepositions with place names which involve the village classifier. In the upper part of the table, those place names are featured whose occurrences are restricted to combinations with

two or only one of the prepositions. Empty cells are additionally marked by grey shading.

type	*ġo*	*ġewwa*	*fi*	sum
Ħas-Serħ		4		4
Ħal Tartani		6		6
Ħal Ferħ		2	38	40
Ħal Mula		1	12	13
Ħas-Saptan		1	6	7
Ħad-Dingli	4	27		31
Ħal Balzan	2	24	657	683
Ħal Far	12	77	1,138	1,227
Ħal Farruġ	16	17	327	360
Ħal Kirkop	5	37	759	801
Ħal Luqa	19	61	1,205	1,285
Ħal Qormi	40	142	3,440	3,622
Ħal Safi	4	34	749	787
Ħal-Saflieni	1	1	5	7
Ħal Tarxien	10	71	1,072	1,153
Ħal Lija	2	16	459	477
Ħaż-Żabbar	22	115	336	473
totals	137	636	10,203	10,976

Table 11: Place names with village classifier in $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$ (tokens)

The most striking aspect about Table 11 is the omnipresence of *ġewwa*. *Ġo* and *fi* are excluded from combinations with five and three place names, respectively. In contrast, *ġewwa* is attested throughout the above list. It almost goes without saying that wherever *fi* is attested it yields the highest token frequency of the three prepositions. *Fi* accounts for 93% of the total number of tokens. As to the number of types, however, *ġewwa* outranks both *fi* and *ġo*. In two cases, *ġewwa* is the only realized option.

On top of that, *ġewwa* is also the uncontested number one if *Ħad-Dingli* occupies the TOP-slot which is a remarkable fact because *fi* fails to show up as a contender. This is certainly an effect of the corpus since the reviewers argue that native speakers have no problem producing, e.g., *f'Ħad-Dingli* 'in Ħad-Dingli'. The prominent position of *ġewwa* is further strengthened by the data in Table 12.

type	***ġo***	***ġewwa***	***fi***	**sum**
Raħal Ġdid	9	120	1,857	1,986
raħal ta' Kerċem		1	9	10
raħal ta' Wied il-Għajn		2	8	10
raħal tal-Iklin		2		2
totals	9	125	1,874	2,008

Table 12: Complex expressions involving *raħal* 'village' (types/tokens)

There are types of expressions which contain the common noun *raħal* 'village' from which the village classifier *Ħal* is derived. In the case of *Raħal Ġdid*, *raħal* forms part of the place name itself. In the remaining three types, *raħal* is exterior to the place name, which means that the constructions with *raħal ta'* 'village of' do not fall into the class of genuine place names. All four types are combinable with *ġewwa*. The place name *Raħal Ġdid* is the only type to combine with all three prepositions. Table 12 reflects the same patterns as Table 11.

We summarize this section as follows. In terms of type frequency, *ġewwa* is the best candidate for combinations with complex place names, especially if these place names also meet the criterion of familiarity. *Fi* is not as sensitive as *ġewwa* to any of these parameters. *Ġo*, in contrast, seems to behave like *ġewwa*, at least with respect to the role of familiarity.

3.4 *Wasal* + [PREP$_{PLACE/GOAL}$ TOP]$_{pp}$

Owing to the relatively low frequency of *ġo* in [PREP$_{PLACE/GOAL}$ TOP]$_{PP}$, we start the final empirical section with an overview of *ġo*-PPs with place names as complement as they are found in Attard (1999). We have opted for this particular text not only because it forms part of the *Korpus Malti 3.0* (academic section) but also because of the striking variation in the use of the relevant constructions. The full range of this variation has to be studied in-depth in the future. In this section, we only touch upon the use of [PREP$_{PLACE/GOAL}$ TOP]$_{PP}$ as complements of one particular motion verb, namely *wasal* 'arrive'.

There are altogether fifty-nine *ġo*-PPs distributed over fifty-eight sentences in our sample text. The examples, together with their immediate syntactic context, are presented in the appendix. The bulk of the *ġo*-PPs form part of descriptions of static situations. Places are identified where something is located or something happened. Existential predicates are numerous. There are also NP-internal *ġo*-PPs which function as prepositional attribute of a head noun. The spatial relation is that of Place, with the sole exception of #47, which we present as (5) in this section.

(5) *ġo* (Attard 1999: 123)

F-l-1947	*wasal*	***[ġo***	***Cardiff]***	*Patri Hugh Attard…*
in-DEF-1947	arrive	**[in**	**Cardiff]**	Father Hugh Attard

'In 1947 Father Hugh Attard arrived **[in Cardiff]**…'

This example is exceptional and therefore interesting because occasionally *wasal* may take a *fi*-PP as well, as can be seen in (6).

(6) *fi* (Attard 1999: 59)

Mons. Caruana	*kien*	*minn*	*Tas-Sliema*	*u*	*wasal*
Mons. Caruana	be.PFV	from	Tas-Sliema	and	arrive
[f-l-Istati Uniti]	*f-l-1910.*				
[in-DEF-US]	in-DEF-1910				

'Mons. Caruana was from Tas-Sliema and he arrived **[in the United States]** in 1910.'

Normally however – and not only in our sample text – *wasal* prefers bare place names as complements over fully-blown PPs. Example (7) is typical for this preference.

(7) Bare place name (Attard 1999: 6)

Meta	*L'Isle Adam*	*wasal*	***[Ø***	***Malta]***	*f-l-1530…*
when	L'Isle Adam	arrive	**[Ø**	**Malta]**	in-DEF-1530

'When L'Isle Adam arrived **[in Malta]** in 1530…'

Bare place names are not the exclusive privilege of *wasal*. In point of fact, the situation is paralleled by other highly frequent motion verbs, such as *mar* 'go' and *daħal* 'enter'. In Attard (1999) there are 269 tokens of motion verbs combining with a place name as Ground. These motion events are a clear minority in comparison to the 660 static spatial situations described in the same source. Of the 269 dynamic situations, 202 (= 75%) go to the credit of direct combinations of motion verb and bare place names. Bare place names are involved in only 16% of all static spatial situations. This is the domain of *fi* which accounts for 71% of all static situations. *Ġo* and *ġewwa* are attested infrequently but seem to display a preference similar to that of *fi* as can be gathered from Diagram 8.

It is tempting to interpret Diagram 8 as a piece of evidence for the existence of a relatively clear-cut division in two. On the one side, there is absence of material encoding of the spatial relations with a leaning towards Goal. On the opposing side, we find the phonologically realized prepositions, all of which tend to receive a static reading unless further information is available. This in turn is suggestive of the almost full synonymy of *ġewwa*, *ġo*, and *fi* in the realm of Place. The above (subtle) differences between the prepositions notwithstanding, their functional domains overlap considerably. Thus, there is no absolutely free variation but we are almost there, in a manner of speaking.

4. Conclusions

The exact nature of the relationship between *fi*, *ġewwa*, and *ġo* has been shown to be hard to pinpoint. The quantitative dominance of *fi* was easy to establish as to the token frequency. The picture is not as straightforward as that if we look at type frequency, where *ġewwa* has its stronghold. *Ġo* can compete with neither of the other two prepositions in terms of the overall frequencies. However, it would be wrong to sweepingly assume that *ġo* is a negligible entity in the prepositional system of Maltese. In connection with the construction type $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$, *ġo* and *ġewwa* display a relatively high degree of sensitivity to criteria such as the familiarity of the place name, the ontological class of the geo-object, and the complexity of the place name. This sensitivity is only relative because what results from it can only be captured with the term slight preference – a preference which is manifest more on the level of type frequencies than on that of token frequencies. None of the preferences discussed in Sections 3.1–3.3 is strong. The choice of preposition depends on two or more factors conspiring, or so it seems. We have not been able to discover any strict rules which determine what option is the best

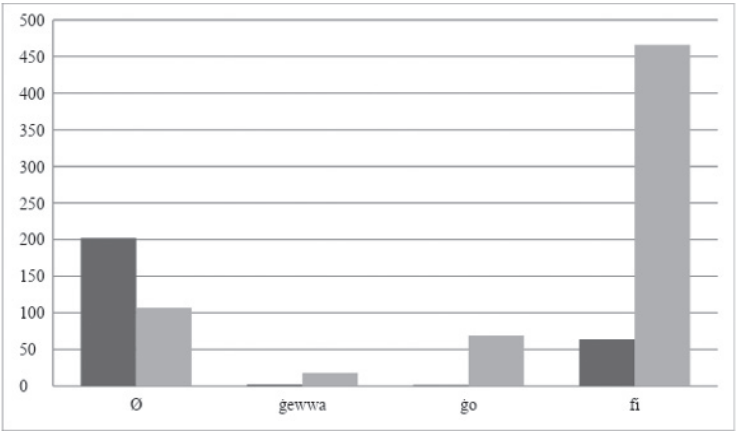


Diagram 8: Coding strategies and dynamic vs static situations in Attard (1999)

in a given context. There is one exception though, namely, the underrepresentation of motion events in the situations encoded by *fi* and especially *ġewwa* and *ġo*. This is where the fourth option – zero-marking – enters the scene. At this point of our investigation, it seems that there is a bipartition of zero-marking vs overt marking.

The project still has a long way to go. To prove that STG is well-established in Maltese, we need to compare our findings in connection to $[\text{PREP}_{\text{PLACE/GOAL}}\ \text{TOP}]_{\text{PP}}$ with constructions that involve a common noun as Ground. Moreover, the data from the *Korpus Malti 3.0* have to be analyzed further and thoroughly to make it possible for us to find out about potential syntactic factors which have a say in the choice of preposition. It is certainly insufficient to take account of only one particular motion verb. Many kinds of predicates must be analyzed before we can declare the project closed. It is hoped that this study can convince scholars of Maltese, experts of STG, and those who are interested in adpositions that it is fruitful to inquire into the grammar of Maltese prepositions.

Acknowledgments

This study forms part of the research project *Präpositionen und ihre Grammatik im Maltesischen / Prepositions and their grammar in Maltese* (grant number STO186/22-1) funded by the *Deutsche Forschungsgemeinschaft (DFG)*. The paper elaborates upon preliminary ideas which were presented as a talk entitled *Different Maltese ways of being in Sydney (and elsewhere). On the formal variation of expressing Place with toponyms* on occasion of the *7th Lingwistika Maltija* (Kraków, 10–12 July, 2019). We are grateful to our audience for their helpful comments. We are also indebted to Ray Fabri and Michael Spagnol as well as our anonymous reviewers for commenting on the draft version of this study. We assume full responsibility for all remaining errors.

Abbreviations

3	third person	PPTCPL	passive participle
DEF	definite article	PREP	preposition
IPFV	imperfective	PL	plural
PFV	perfective	STG	Special Toponymic Grammar
PP	prepositional phrase	TOP	toponym
NP	noun phrase		

Appendix

[page numbers (= _X_) refer to Attard (1999)]

1. _11_ *...biex jara liema kienu l-prospetti għal pjan ta' emigrazzjoni tal-Maltin **[ġo Ċipru]***. '…to see what the prospects were for a plan of Maltese emigration **[to Cyprus]**.'
2. _11_ *...bis-saħħa ta' element Malti **[ġo Ċipru]** l-Imperu jsaħħaħ l-interessi tiegħu.* '…with a Maltese element **[in Cyprus]** the Empire will strengthen its interests.'
3. _12_ *...sabiex jinbena raħal Malti **[ġo Ċipru]***. '…so that a Maltese village will be built **[in Cyprus]**.'
4. _17_ *Kien hemm gruppi ta' Maltin ġol-belt kapitali Tuneż kif ukoll **[ġo Susa, Monastir, Medhia, Sfax]** u fil-gżira ta' Ġerba.* 'There were groups of Maltese in the capital city Tunis as well as **[in Susa, Monastir, Medhia, Sfax]** and on the island Djerba.'
5. _21_ ***[Ġo Chambray]** kien hemm ħamsin Malti.* '**[In Chambray]** there were fifty Maltese.'
6. _24_ *Kellu interessi navali kemm **[ġo Port Colbourne]** kif ukoll **[ġo St. Catherine's]**...* 'He had naval interests not only **[in Port Colbourne]** but also **[in St. Catherine's]**…'
7. _38_ *...li kien il-konslu tar-Renju Unit **[ġo Sao Paolo]**...* '…who was the consul of the UK **[in Sao Paolo]**…'
8. _40_ *...kienu ġa stabbiliti **[ġo Kalifornja]***. '…they were established already **[in California]**.'
9. _41_ *Din il-knisja kienet għal bosta snin iċ-ċentru tal-ħajja Maltija **[ġo San Francisco]***. 'This church was for many years the center of Maltese life **[in San Francisco]**.'
10. _42_ *...kienu nġabru **[ġo Toronto]***... '…they had gathered **[in Toronto]**…'
11. _46_ *...Ġużeppi sar magħruf sewwa **[ġo Blacktown]***. '…Ġużeppi became very famous **[in Blacktown]**.'
12. _47_ *...**[ġo Tamworth]** bena post kbir għat-trobbija tagħhom.* '…**[in Tamworth]** he built a big house for their upbringing.'
13. _55_ *...xi 2000 Malti jgħixu u jaħdmu **[ġo Margo]***. '…some 2,000 Maltese were living and working **[in Margo]**.'
14. _55_ ***[Ġo Korfù]** u l-gżejjer l-oħra fil-qrib nibtu xi kommunitajiet żgħar ta' Maltin…* '**[On Corfu]** and other islands nearby some small Maltese communities emerged…'

15. _55_ *...grupp ta' sorijiet Maltin fetħu dar **[ġo Korfù]*** ... '…a group of Maltese nuns opened a house **[on Corfu]**.'
16. _56_ ***[Ġo Konstantinopli]*** *l-għadd tal-Maltin ġie li wasal għal tliet elef ruħ.* '**[In Constantinople]** the number of Maltese reached 3,000 souls.'
17. _56_ *B'hekk il-preżenza Maltija **[ġo Konstantinopli u Smyrna]** tista' tgħid li ġiet fix-xejn.* 'In this way the Maltese presence **[in Constantinople and Smyrna]** came to an end, one may say.'
18. _59_ *Ħadem fost il-Maltin u t-Taljani kemm fi New York kif ukoll **[ġo Detroit]***. 'He worked among the Maltese and Italians in New York as well as **[in Detroit]**.'
19. _62_ *Miet **[ġo Huntingdon]*** ... 'He died **[in Huntingdon]**…'
20. _63_ *Id-dimostrazzjoni kienet saret kontra s-sensji minn ma' Ford **[ġo Dearborn]***. The demonstrations had taken place against the discharges from Ford **[in Dearborn]**.'
21. _64_ *...kien r-rappreżentant tal-provinċja ta' Quebec **[ġo Londra]***. '…he was the representative of the province Quebec **[in London]**.'
22. _65_ *Il-preżenza Maltija **[ġo Toronto]** kienet ilha tinħass sa mis-seklu dsatax*... 'The Maltese presence **[in Toronto]** had been visible since the 19th century…'
23. _65_ *...kienu qed jgħixu **[ġo Toronto]***. '…they were living **[in Toronto]**.'
24. _65_ *...l-ħajja tal-Maltin **[ġo Toronto]** ma kinitx waħda komda.* '…the life of the Maltese **[in Toronto]** was not an easy one.'
25. _73_ *...in-numru tal-Maltin **[ġo Victoria]** kien ta' madwar erba' mija.* '…the number of Maltese **[in Victoria]** was about 400.'
26. _73_ ***[Ġo Western Australia]*** *kien hemm xi mitejn Malti*… '**[In Western Australia]** there were some 200 Maltese…'
27. _74_ ***[Ġo Mackay]*** *il-familja Busuttin kienet magħrufa sewwa*… '**[In Mackay]** the family Busuttin was very well-known…'
28. _76_ *Meta ltaqa' l-Kungress Ewkaristiku Internazzjonali **[ġo Sydney]** fl-1928*... 'When the International Eucharistic Congress met **[in Sydney]** in 1928…'
29. _76_ *Skond Parnis fl-1929 **[ġo Sydney]** kien hemm madwar erba' mija u ħamsin Malti*… 'According to Parnis, there were about 450 Maltese [in Sydney] in 1929…'
30. _95_ ***[Ġo Aden]*** *sirna nafu li f'Malta kienet waslet l-aħbar*… '**[In Aden]** we came to know that the news had arrived in Malta…'
31. _98_ *...Alfred u Aida ġew mistiedna għal pranzu **[ġo Canberra]** mill-Ministru*… '…Alfred and Aida were invited for dinner **[in Canberra]** by the Minister…'
32. _100_ *...**[ġo N.S.W.]** familja ta' ħamsa kienet qed tħallas £1 fil-ġimgħa*… '… **[in New South Wales]** a family of five was paying £1 per week…'
33. _100_ *Fil-fatt l-iskema bdiet taħdem **[ġo Sydney]** fit-8 ta' Awissu 1967.* 'In fact the scheme began to work **[in Sydney]** on 8 August, 1967.'
34. _104_ *Il-President Kennedy ġie maqtul **[ġo Texas]** fit-22 ta' Novembru 1963.* 'President Kennedy was murdered **[in Texas]** on 22 November, 1963.'
35. _109_ *...il-Kummissarju tar-Renju Unit **[ġo Ottawa]** għarraf lil Mr Jollife*... '…the Commissioner of the UK **[in Ottawa]** explained to Mr Jollife…'

36. _110_ *...fl-1947 kien ġa attiv ħafna **[ġo Ottawa]*...* ‘…in 1947 he was already very active **[in Ottawa]**…
37. _112_ *...il-ħajja kemm **[ġo Fingal]** kif ukoll **[ġo Ajax]** kienet xi ftit reġimentata*... ‘…life was a bit strict **[at Fingal]** as well as **[at Ajax]**…’
38. _113_ *...dawk li baqgħu **[ġo Ajax]*...* ‘…those who remained **[at Ajax]**…’
39. _114_ *Fil-fatt x'kienet il-verità dwar il-ħajja **[ġo Ajax]**?* ‘What in fact was the truth about life **[at Ajax]**?’
40. _114_ ***[Ġo Ajax]** kien hemm kappella, skola, u sptar.* ‘**[At Ajax]** there was a chapel, a school, and a hospital.’
41. _115_ *Minkejja dan is-saram li nqala‘ **[ġo Ajax]*...* ‘In spite of the confusion which happened **[at Ajax]**…’
42. _117_ *Kien hemm xi ftit **[ġo British Columbia, Manitoba, u Quebec]*...* ‘There were some **[in British Columbia, Manitoba, and Quebec]**…’
43. _118_ *…ġie maħtur bħala kummissarju tal-Gvern Malti **[ġo Londra]*...* ‘…he was elected commissioner of the Maltese government **[in London]**…’
44. _118_ *...kien hemm xi ftit Maltin **[ġo Chatham u Portsmouth]**.* ‘…there were some Maltese **[in Chatham and Portsmouth]**.’
45. _119_ *...għal xi Maltin li għamlu għajb **[ġo Londra]** lill-isem ta‘ ġensna.* ‘…because of some Maltese who dishonored the name of our people **[in London]**.’
46. _123_ *Fl-1969 kien hemm **[ġo Portsmouth]** madwar mija u sebgħin minn dawn in-nisa*... ‘In 1969 there were about 170 of these women **[in Portsmouth]**… ’
47. _123_ *Fl-1947 wasal **[ġo Cardiff]** Patri Hugh Attard*... ‘In 1947 Father Hugh Attard arrived **[in Cardiff]**…’
48. _123_ *...dawk li kienu jinsabu **[ġo Barry u Newport]**.* ‘…those who found themselves **[in Barry and Newport]**.’
49. _124_ *...mitt raġel ġew nominati għal xogħol **[ġo Canberra]*...* ‘…100 men were nominated for work **[in Canberra]**.’
50. _124_ *Imma **[ġo Canberra]** nqala’ saram*... ‘But **[in Canberra]** there was confusion…’
51. _128_ *Kellu wkoll laqgħa ma’ nies prominenti **[ġo Victoria]*...* ‘He also had a meeting with prominent people **[in Victoria]**…’
52. _143_ *...is-snin li għaddew **[ġo Tardun]**.* ‘…the years they passed **[in Tardun]**.’
53. _143_ ***[Ġo Clontarf]** AM kien ma’ grupp ta‘ mitejn u erbgħin tifel*... ‘**[At Clontarf]** AM was together with a group of 240 boys…’
54. _146_ *...jattendu xi lezzjonijiet taħt Sr Francis Margaret **[ġo Ħal Balzan]**.* ‘…they will attend classes under Sr Francis Margaret **[in Ħal Balzan]**.’
55. _148_ *Meta waslu sabu xogħol **[ġo Coburg]**.* ‘When they arrived they found work **[in Coburg]**.’
56. _150_ *Fl-1971 is-Sur Forace laħaq Kummissarju Għoli ta’ Malta **[ġo Canberra]**.* ‘In 1971 Sur Forace became High Commissioner of Malta **[in Canberra]**.’
57. _159_ *Aktar tard dehru ġurnali Maltin **[ġo Detroit]*...* ‘Later Maltese journals were published **[in Detroit]**…’
58. _161_ *Iċ-ċeremonja saret **[ġo Sydney]*...* ‘The ceremony took place **[in Sydney]**…’

References

Anderson, John M. (2007) *The grammar of names.* Oxford: Oxford University Press.

Aquilina, Joseph (1987) *Maltese-English dictionary- vol. 1: A-L.* Malta: Midsea.

Attard, Lawrence E. (1999) *L-emigrazzjoni Maltija. Is-seklu dsatax u għoxrin.* Il-Pjetà: Pubblikazzjonijiet Indipendenza.

Borg, Albert & Azzopardi-Alexander, Marie (1997) *Maltese. Lingua descriptive grammar.* London & New York: Routledge.

Nübling, Damaris, Fahlbusch, Fabian & Heuser, Rita (2015) *Namen. Eine Einführung in die Onomastik*. Tübingen: Narr.

Nyström, Staffan (2016) Names and meaning. In: Hough, Carol (ed.), *The Oxford handbook of names and naming.* Oxford: Oxford University Press, 39–51.

Saari, Rami (2003) *Milot Hayaxas Hamalteziyot.* Jerusalem: Carmel.

Schmidt, Emeli & Vorholt, Maike & Witt, Nele (2020) Form and behavior of Maltese prepositions – A usage-based approach. In: Čéplö, Slavomír & Drobný, Jaroslav (eds.), *Maltese linguistics on the Danube.* Berlin & Boston: De Gruyter, 241–270.

Stolz, Thomas & Ahrens, Carolin (2017) On prepositional ellipsis and the factors which block its application in Maltese. In: Saade, Benjamin & Tosco, Mauro (eds.), *Advances in Maltese linguistics* (= Studia Typologica 21). Berlin & Boston: De Gruyter Mouton, 121–147.

Stolz, Thomas, Lestrade, Sander & Stolz, Christel (2014) *The crosslinguistics of zero-marking of spatial relations.* Berlin & Boston: De Gruyter Mouton.

Stolz, Thomas, Levkovych, Nataliya & Urdze, Aina (2017a) Die Grammatik der Toponyme als typologisches Forschungsfeld. *Linguistische Berichte*. Sonderheft 23, 121–146.

Stolz, Thomas, Levkovych, Nataliya & Urdze, Aina (2017b) When zero is just enough… In support of a special toponymic grammar in Maltese. *Folia Linguistica* 51(2), 453–482.

Stolz, Thomas, Levkovych, Nataliya & Urdze, Aina (2018) La morfosintassi dei toponimi in prospettiva tipologica. In: Brincat, Giuseppe & Caruana, Sandro (eds.), *Tipologia e 'dintorni': il metodo tipologico alla intersezione di piani d'analisi*. Roma: Bulzoni, 307–324.

Stolz, Thomas & Levkovych, Nataliya (2019a) Toponomastics meets linguistic typology: Glimpses of special toponymic grammar from Aromanian and sundry languages. *Onomastica Uralica* 11, 43–61.

Stolz, Thomas & Levkovych, Nataliya (2019b) Absence of material exponence. *Language typology and universals – STUF* 72(3), 373–400.

Stolz, Thomas & Levkovych, Nataliya (2020) From variation to the grammar of Maltese prepositions – first steps. In: Čéplö, Slavomír & Drobný, Jaroslav (eds.), *Maltese linguistics on the Danube.* Berlin & Boston: De Gruyter, 199–240.