A Head-Driven Phrase Structure Grammar for Maltese

Stefan Müller
Freie Universität Berlin

February 20, 2009

Abstract

This paper describes a grammar fragment of Maltese that is implemented in the framework of Head-Driven Phrase Structure Grammar.

1 Introduction

It was Utz Maas who first aroused my interest in the Maltese language and asked me to discuss with the Osnabrück group how HPSG could be applied to Maltese. I started to read the sparse literature on Maltese and discovered Ray Fabri's book on Maltese grammar. Fabri's analyses are coached in the framework of Lexical Decomposition Grammar (Wunderlich, 1997), which shares a lot of basic assumptions with HPSG, so that this book was ideal as a background for my project. I quickly discovered that the interactions of phenomena are quite complex and that a computer implementation of my theoretical ideas would be helpful. I started to translate Fabri's analyses into HPSG analyses, and extended, adapted or changed them when I found this necessary. HPSG differs from LDG in allowing types as values of features instead of binary values only. This makes it possible to use types for classifying linguistic objects and makes fewer features necessary. Apart from this formal difference, I refrained from using defaults. The HPSG grammar uses Minimal Recursion Semantics (Copestake, Flickinger, Pollard and Sag, 2005) as the underlying semantic formalism and is implemented in the TRALE system (Meurers, Penn and Richter, 2002; Penn, 2004).

2 Background and Basic Assumptions

Head-Driven Phrase Structure Grammar (HPSG) is a grammatical framework that was developed in the 80s by Pollard and Sag (Pollard and Sag, 1987, 1994). Since then more than 1,000 papers on HPSG were published and a large number of languages were dealt

[†]This work was presented in 2007 in Osnabrück at the Chair for General and Germanic Linguistics and at the first international conference on Maltese Linguistics in 2007 in Bremen. The section concerning constituent order was presented at the 2008 HPSG conference in Kyoto.

I thank Utz Maas for the invitation to Osnabrück and Felix Bildhauer, Ray Fabri, Utz Maas, John Peterson, and Thomas Stolz for discussion and Felix Bildhauer, John Peterson, Frank Richter, Thomas Stolz, and an anonymous reviewer for comments on an earlier draft of this paper.

with theoretically and for many languages computationally processable grammars were implemented.

Typed feature structures are used to model linguistic objects. Linguists use typed feature descriptions to describe stems, words, lexical rules, and phrases. In what follows I will sketch some basic assumptions. Due to space limitations this part will be very brief. The interested reader is referred to Pollard and Sag, 1994 and to the textbooks of Borsley, 1999, Sag, Wasow and Bender, 2003, and Müller, 2008.

2.1 Valency

Every lexical item comes with a list that describes the arguments that have to be combined with it in order to yield a complete phrase. (1) gives some examples:

(1)

verb	argument structure
ġie ('to come')	⟨ NP[*str*] ⟩
kiteb ('to write')	⟨ NP[*str*], NP[*str*] ⟩

ġie requires one NP with structural case and *kiteb* two. Case in verbal environments is assigned by the Case Principle as follows:[1]

Principle 1 (Case Principle)

- *In a list that contains both subjects and complements of a verbal head, the first element with structural case is assigned nominative case unless it is raised to a dominating head.*
- *All other elements of this list with structural case are assigned accusative case.*

This Case Principle is very similar to the one that was suggested by Yip, Maling and Jackendoff (1987) and therefore also accounts for the case systems of a variety of languages that were discussed by these authors, including the complicated case system of Icelandic. The case principle above is also assumed in HPSG grammars of German (Meurers, 1999; Müller, 2008), Finnish (Przepiórkowski, 1999a), and Hindi (Müller and Vasishth, In Preparation).

The elements in the valency list are linked to argument slots in semantic representations. (2) shows an example entry for *ġie*:

(2)
$$\left[\begin{array}{ll}
\textsc{phon} & \langle\, \textit{ġie} \,\rangle \\
\textsc{synsem} & \left[\begin{array}{ll}
\textsc{loc} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}
\textsc{head} & \textit{verb} \\
\textsc{arg-st} & \left\langle\, \textsc{np}[\textit{str}]\boxed{1} \,\right\rangle
\end{array}\right] \\
\textsc{cont} & \left[\textsc{ind}\ \boxed{2}\right]
\end{array}\right] \\
\textsc{nonloc} & \ldots
\end{array}\right] \\
\textsc{rels} & \left\langle \left[\begin{array}{ll}
\textsc{pred} & \textit{come} \\
\textsc{arg0} & \boxed{2}\,\textit{event} \\
\textsc{arg1} & \boxed{1} \\
\textit{arg1-relation} &
\end{array}\right] \right\rangle
\end{array}\right]$$

[1] For a way to formalize principle 1 see Przepiórkowski, 1999b; Meurers, 1999; Meurers, 2000, Chapter 10.4.1.4.

Phonological information is presented under PHON. The value of PHON is a list of phonemes. Instead of a phoneme representation I give an orthographic representation to increase readability.

Syntactic and semantic information that can be selected by other heads is represented under SYNSEM. The value of SYNSEM is structured: LOC contains information that is relevant locally, NONLOC is used to model nonlocal dependencies (see Section 8.2). CAT contains the syntactic category information like part of speech (represented under HEAD) and valence information and CONT contains selectable semantic information. All information that is contained under HEAD will be projected along the head path in a syntactic analysis. This is ensured by the Head Feature Principle, which identifies the HEAD value of the mother with the HEAD value of the head daughter.

Boxed integers are used in HPSG to mark identity, so the referential index of the NP argument of *ġie* is linked to the ARG1 of the *come* relation. The relations that are contributed by a word or by a phrase are represented as the value of the RELS feature. I use Minimal Recursion Semantics (MRS) for the representation of the semantic contribution (Copestake et al., 2005), but the feature geometry differs from the one assumed in standard MRS in that the relations are not represented under the path SYNSEM|LOC|CONT but at the top level of the feature structure. Since the information that can be selected by other heads is the information grouped under SYNSEM (syntax-semantics), this feature geometry restricts the selectable information to those parts of the semantic representation that have to be accessible, that is the semantic variables that are represented under IND (index). See also Sailer, 2004 on the locality of selection with regard to semantic information.

2.2 Basic Constituent Structure and Complementation

I follow Fabri (1993) in assuming that Maltese is a language with a verb phrase in VO order. The subject may be realized either to the left or to the right of this VP (p. 142). Pollard and Sag (1994, Chapter 9) and Sag, Wasow and Bender (2003, Chapter 4) suggest using two features for the representation of arguments for a configurational language like English: the SPR feature to represent the specifier/subject and the COMPS feature to represent all other arguments. I use the more traditional SUBCAT feature instead of COMPS. The reason is that I follow Kiss (1992; 1995) in assuming that the subject of finite verbs in German is represented on the same list as the other arguments of the verb. This accounts for the fact that German allows the subject to be serialized freely with regard to the other arguments. Languages differ in the way they map the arguments from ARG-ST to SPR and SUBCAT. While in German all arguments of finite verbs are mapped to SUBCAT, the arguments of Maltese and English are split: The subject is represented as the only member of the SPR list and the other arguments are mapped to SUBCAT.

In the analysis presented below a verb is combined with the element on SUBCAT to form a VP and in the next step this VP is combined with its specifier to form a complete sentence.

(3) L-iskrivan kiteb l-ittr-a.[2]
def-employee.3msg wrote.3msg def-letter-fsg
'The employee wrote the letter.'

[2] Fabri, 1993, p. 130.

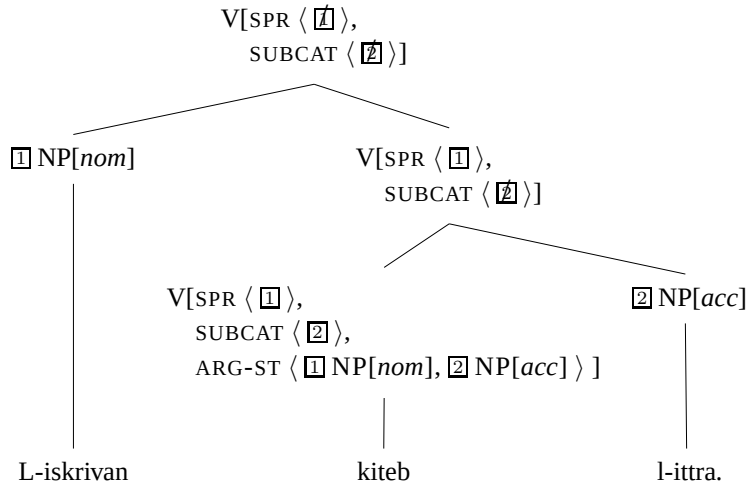


Figure 1: Basic Clause Structure

The analysis in Figure 1 differs from standard HPSG in not canceling off the elements in the valence representation after combination: in standard HPSG the SUBCAT list of the VP would be the empty list, but in Figure 1 this SUBCAT list still contains the element. Note though that this element is marked as saturated (see also Higginbotham, 1985, p. 560 for a similar proposal in a different framework). Meurers (1999) coined the term *spirit* for realized arguments that still hang around. While this approach to saturation may not be necessary for Maltese, it is necessary to get the case assignment facts in German right and if one assumes that basic principles of grammar are the same across languages, valence representations and saturation of arguments should be similar across languages.[3]

Syntax trees are just a means of visualization in HPSG. All linguistic objects (roots, stems, words, clitics, phrase schemata, and lexical rules) are modeled by typed feature structures. The fact that the phrase *kiteb littra* consists of two daughters, a head daughter and a non-head daughter, can be represented as in (4):

(4)
$$\begin{bmatrix} \text{PHON} & \langle\, \textit{kiteb}, \textit{littra} \,\rangle \\ \text{HEAD-DTR} & \left[\text{PHON} \;\langle\, \textit{kiteb} \,\rangle\right] \\ \text{NON-HEAD-DTRS} & \left\langle \left[\text{PHON} \;\langle\, \textit{littra} \,\rangle\right] \right\rangle \end{bmatrix}$$

Basically this is a one by one description of the tree constituents. The value of NON-HEAD-DTRS is a list since this allows for structures with several non-head daughters. An example of a construction that can be analyzed with two non-head daughters are relative clauses, which can be described as binary headless structures (Müller, 1999, Chapter 10) that provide a nominal semantics despite their being headed by a verb internally.

The linguistic object that models *kiteb littra* is of the type *head-argument-phrase* and is constrained by the following constraint:

[3]Note that I do not make any claims about innateness here.

Schema 1 (Head Argument Schema (binary branching))
head-argument-phrase →

$$\begin{bmatrix} \text{SUBCAT } \boxed{1} \oplus \langle \not\boxed{2} \rangle \oplus \boxed{3} \\ \text{HEAD-DTR}|\text{SUBCAT } \boxed{1} \oplus \langle \boxed{2} \rangle \oplus \boxed{3} \textit{ list of spirits} \\ \text{NON-HEAD-DTRS } \left\langle \left[\text{SYNSEM } \boxed{2}\right] \right\rangle \end{bmatrix}$$

This schema requires that the head daughter has a SUBCAT list that can be split into three parts: a possibly empty first part (1), a list that contains an unrealized argument (2), and a possibly empty list of arguments that are already realized (3). The SUBCAT list of the mother differs from the SUBCAT list of the head daughter only in that the element 2 is marked as realized. The description of the argument in the valence representation of the head has to match the properties of the element in the list of non-head daughters.

3 Morphology

The grammar fragment covers basic forms of inflection (inflection for masculine and feminine gender, verbal inflection), definiteness marking, and cliticization.

3.1 Inflection

The Maltese lexicon has to contain root forms specifying consonant patterns and information about vowels that have to be inserted to derive stems. (5) shows some examples:

(5)	citation form	root	vowel melody	imperative
	ħolom	ħlm	o	oħlom ‘dream’
	ħadem	ħdm	ae	aħdem ‘work’
	niżel	nżl	ie	inżel ‘go down’
	seraq	srq	ea	israq ‘steal’

The relation between roots and stems can be implemented via a lexical rule. The lexical rule takes as input the lexical description of a verb and inserts the vowels in the respective position in the phonological representation of the output. Other lexical rules relate such stem representations to fully inflected words that can be used in syntax.

Nouns can be grouped into nouns with a sound (6) and into nouns with a so-called broken plural (7):

(6) a. art
country

b. art-ijet
countries

(7) a. bolla
stamp

b. bolol
stamps

c. tabib
doctor

d. tobba
doctors

While the stems of regular nouns can be related to words by the simple concatenation of phonological material, the forms in (7) require insertion of vowels in consonant

patterns. Both cases can be handled by lexical rules relating lexical representations to fully inflected words.

Usually, adjectives end in *-a*, if they refer to feminine nouns:

(8) a. ktieb sabiħ
book.msg beautiful.msg

b. vedut-a sabiħ-a
view-fsg beautiful-fsg

However, many non-native adjectives do not inflect for gender:

(9) a. raġel intelliġenti
man.msg intelligent
'an intelligent man'

b. mar-a intelliġenti
woman-fsg intelligent
'an intelligent woman'

(10) shows the lexical rule that licences feminine forms:

(10) $$\begin{bmatrix} \text{PHON } \boxed{1} \\ \text{SYNSEM|LOC|CONT|IND|GEN } \textit{fem} \\ \textit{native\_stem} \end{bmatrix} \mapsto \begin{bmatrix} \text{PHON } \boxed{1} \oplus \langle\, a \,\rangle \\ \textit{word} \end{bmatrix}$$

Gender information is part of the referential index that is part of the semantic representation of nouns and adjectives (Pollard and Sag, 1994). The lexical rule applies to a stem with feminin gender as described in the lefthand side of the rule and licences a word, as described in the righthand side of the rule. To take an example the word *sabiħa* is licensed by the stem *sabiħ*.

By convention, values in the input that are not changed explicitly in the output are taken over from the input. There is also a parallel lexical rule for the masculine form that differs only by not adding any phonological material. The latter lexical rule licences forms like *sabiħ* in (8a).

3.2 Definiteness Marking

Definiteness is marked with an /l/ at nouns and adjectives in Maltese. (11) gives an example:[4]

(11) l-ktieb
def-book

If the noun starts with one of the coronals /d/, /t/, /s/, /z/, /ʃ/, /ts/, /tʃ/, /n/, or /r/ the /l/ is assimilated. (12) gives an example:

(12) r-raġel
def-man

The only exception is the coronal /dʒ/, which is exempt from assimilation.

Inner epenthesis can be observed if the word starts with /s/ or /ʃ/ followed by a consonant:

[4](11), (12) and (13b) are underlying forms. If the definite form of *book* is used in isolation, an /i/ has to be added. See below.

(13) a. skola
school

b. l-iskola
def-school

If inner epenthesis applies, it prevents assimilation.

Outer epenthesis happens if the word to which the definiteness marker attaches starts with a consonant and it is not the case that there is a preceding word that ends in a vowel (Fabri, 1993, p. 41).

(14) a. Ta-ni l-ktieb.
gave-me def-book
'He gave me the book.'

b. Ta-k il-ktieb.
gave-you def-book

The formulation of the conditions for outer epenthesis seems complicated, but it gets the facts right at the beginning of sentences.

The following lexical rule accounts for the definiteness marking:

(15)
$$\begin{bmatrix} \text{PHON } \boxed{1} \\ \text{SYNSEM}|\text{LOC}|\text{CAT} \begin{bmatrix} \text{HEAD} & \textit{nominal} \\ \text{DEF} & - \\ \text{MARKING} & \textit{unmarked} \end{bmatrix} \\ \textit{word} \end{bmatrix} \mapsto \begin{bmatrix} \text{PHON f}(\boxed{1}) \\ \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{DEF} + \end{bmatrix}$$

f is a function that respects the phonological restrictions stated above and attaches an /l/ or a coronal. Lexical entries of nominal elements (adjectives and nouns) that are specified in the lexicon have the DEF value $-$. Definite forms are licenced by the lexical rule that takes these elements as input. For instance, the rule may apply to *ktieb* and licence *l-ktieb*. That is the rule above should be understood as a statement saying: whenever we have something that matches the lefthand side of the rule, there is also something that matches the righthand side.

The value MARKING is important to get the facts about *lil* marking right: The definiteness marker does not attach to *lil* marked nouns. See Section 6 on *lil* marking.

4 Noun Phrases

Noun phrases can consist of demonstratives and a common noun:

(16) dan il-ġurdien
this def-mouse

If a noun is used together with a demonstrative, it has to be marked as definite. Human nouns can be used together with the numeral *wieħed* (msg) or *waħda* (fsg). Usually these numerals are realized postnominally, but in the cases under discussion they appear prenominally. Fabri (2001, p. 155) treats these elements as indefinite determiners:

(17) waħd-a mara
INDEF-fsg woman
'a woman'

I follow Fabri (2001) in assuming that the demonstrative and the indefinite article are determiners in nominal structures. However, I do not assume a DP structure but treat the noun as the head in nominal structures. Nouns select a determiner via the SPR feature.[5] Following Pollard and Sag (1994, p. 50), I assume that determiners can select the head they specify as well. This selection is established via a feature SPEC and the SPEC feature principle, which ensures that the SPEC value of a non-head is identified with the SYNSEM value of the head. By this mechanism the demonstrative pronoun can impose the definiteness restriction on the noun it combines with. The demonstrative introduces a general quantifier. The lexical rule for definiteness marking in (15) does not introduce the ι-operator, since if it did, we would bind the same variable by two quantifiers and the formula would be ill-formed. (18) shows the lexical items for the demonstrative pronoun and the noun:

(18) a. *dan* ('this'):

$$\begin{bmatrix} \text{SYNSEM}|\text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{SPEC}|\text{LOC} & \begin{bmatrix} \text{CAT}|\text{DEF} & + \\ \text{CONT}|\text{IND} & \boxed{1} \end{bmatrix} \end{bmatrix} \\ & \textit{det} \\ \text{SPR} & \langle\rangle \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{ARG0} & \boxed{1} \\ \textit{demonstrative\_q} \end{bmatrix}, \begin{bmatrix} \text{ARG1} & \boxed{1} \\ \textit{proximal} \end{bmatrix} \right\rangle \end{bmatrix}$$

b. *il-ġurdien* ('def-mouse'):

$$\begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{noun} \\ \text{DEF} & + \\ \text{SPR} & \langle\, \text{DET} \,\rangle \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{CONT} & \begin{bmatrix} \text{IND} & \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{RELS} & \left\langle \begin{bmatrix} \text{ARG0} & \boxed{1}\, \textit{ref} \\ \textit{mouse} \end{bmatrix} \right\rangle \end{bmatrix}$$

The lexical items are not complete: in order to get the MRS apparatus to work, constraints on possible scopings have to be stated, that is, the fact that the relations contributed by the nominal projection are within the scope of the quantifier has to be represented somehow. Due to space limitations I do not introduce all details concerning MRS here and point the reader to Copestake, Flickinger, Pollard and Sag, 2005 instead.

Figure 2 on the following page shows the analysis of (16).

The mother node is licenced by the following schema for head-specifier phrases:

[5] See van Langendonck, 1994, Hudson, 2004, and Müller, 2008, Chapter 6.6.1 for arguments for the NP analysis.

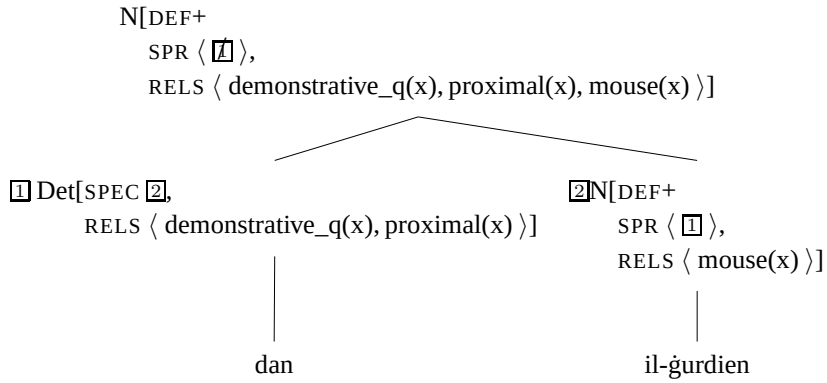


Figure 2: Analysis of *dan il-ġurdien* 'this def-mouse' = 'this mouse'

Schema 2 (Head-Specifier Schema)
head-specifier-phrase →

$$\begin{bmatrix} \text{SYNSEM|LOC|CAT|SPR} \left\langle \boxed{1} \right\rangle \\ \text{HEAD-DTR|SYNSEM|LOC|CAT} \begin{bmatrix} \text{SPR} & \left\langle \boxed{1} \right\rangle \\ \text{SUBCAT} & \textit{list of spirits} \end{bmatrix} \\ \text{NON-HEAD-DTRS} \left\langle \left[\text{SYNSEM } \boxed{1} \right] \right\rangle \end{bmatrix}$$

The head daughter in a head-specifier phrase has to be fully saturated with regard to the SUBCAT elements, that is, the elements in SUBCAT have to be spirits. The single element in the SPR list has to be compatible with the SYNSEM value of the non-head daughter.

Indefinite nouns are used unmarked and without a visible determiner (19a), definite noun phrases have a noun that is marked for definiteness (19b).

(19) a. ktieb
book
'a book'

b. il-ktieb
def-book
'the book'

Adjectives usually follow the noun they refer to and they agree with the noun in number and gender:

(20) a. ktieb sabiħ
book-mas beautiful-mas
'a beautiful book'

b. vedut-a sabiħ-a
view-fsg beautiful-fsg
'a beautiful view'

(21) shows the masculine form of the adjective *sabiħ*.

(21) *sabiħ*

$$\begin{bmatrix} \textsc{phon} & \langle \textit{sabiħ} \rangle \\ \textsc{synsem} & \begin{bmatrix} \textsc{loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{mod} & \overline{\text{N}}[\textit{sg, mas}]\boxed{1} \\ \textsc{pre-modifier} & - \\ \textit{adj} & \end{bmatrix} \\ \textsc{subcat} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \textsc{rels} & \left\langle \begin{bmatrix} \textsc{arg}1 & \boxed{1} \\ \textit{beautiful} & \end{bmatrix} \right\rangle \end{bmatrix}$$

Adjuncts are treated as functors in HPSG: the adjective selects the $\overline{\text{N}}$ that it modifies via the feature MOD (MODIFIED). The adjective specifies the gender and number values of the nominal projection it can modify. This specification ensures the agreement of noun and adjective in gender. See Wechsler and Zlatić, 2003 for a detailed discussion of agreement in HPSG.

All modifiers have a head feature PRE-MODIFIER. The value of this feature is '+', if the modifier precedes the head it attaches to and '−' if it follows the head.

The referential index of the selected nominal projection in (21) is identified with the argument of the *beautiful* relation.

The following schema licences head-adjunct structures:

Schema 3 (Head-Adjunct Schema)
head-adjunct-phrase →

$$\begin{bmatrix} \textsc{head-dtr}|\textsc{synsem} & \boxed{1} \\ \textsc{non-head-dtrs} & \left\langle \begin{bmatrix} \textsc{synsem}|\textsc{loc}|\textsc{cat} & \begin{bmatrix} \textsc{head}|\textsc{mod} & \boxed{1} \\ \textsc{spr} & \textit{list of spirits} \\ \textsc{subcat} & \textit{list of spirits} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

The non-head daughter has to be fully saturated. Its MOD value is identified with the SYNSEM value of the head daughter. This ensures that restrictions that are specified in the lexical entry of a modifier are enforced onto the head in head-adjunct structures. See Figure 3 on page 13 for an analysis of a full NP containing an attributive adjunct.

Nouns and adjectives agree with respect to definiteness in Standard Arabic (Melnik, 2006), but according to Fabri (1993, p. 43; 2001, p. 158) this is not true for Maltese. For Maltese the restriction is: If a noun in an attributive construction is indefinite, the adjective has to be indefinite as well. However, the reverse direction does not hold:

(22) a. il-mara t-twila
def-woman def-tall
b. il-mara twila
def-woman tall
c. il-ktieb il-ġdid
def-book def-new
d. il-ktieb ġdid
def-book new

In addition to the cases with the standard order noun-adjective in (22) there are marked cases in which the adjective appears prenominally. If the NP is definite, the marking appears on the adjective, but not on the noun (see also Kolliakou, 2004, p. 264, p. 273 on non-intersective definite adjectives in Modern Greek).

(23) l-allegat qattiel
def-alleged murderer
'the alledged murderer'

Thus, the NP *l-allegat qattiel* is definite although the noun is not marked as such. I therefore assume that the DEF value of a nominal projection is dependent on the leftmost daughter. In the case of (22) this is the noun, but in (23) it is the adjective.

Fabri (1993) suggested the lexical introduction of the definiteness operator, but revised this analysis in his 2001 paper. I follow him in this analysis and assume that the definiteness affix is just a formal marking on both nouns and adjectives (see the lexical rule in (15) on page 7). The value of the feature DEF is set from $-$ in the input to + in the output of the lexical rule. Determiners are sensitive to the value of DEF of the nominal projection they are combined with. Apart from demonstrative pronouns and the indefinite determiners there are two empty determiners, one with definite and another one with indefinite semantics. The determiner with definite semantics selects via SPEC a nominal projection that has the DEF value + and the indefinite determiner selects one that has the DEF value $-$.

The lexical rule does not introduce semantic material like the ι-operator and therefore can account for the definiteness marking of both adjectives and nouns. If the semantic contribution would be introduced by the lexical rule, the same variable would be bound by two operators in cases like (22a, c), in which both the noun and the adjective are marked for definiteness.

However, the definiteness marking on the adjective in examples like (22a, c) has a semantic effect. According to Fabri (2001, p. 171), (22c) can only be used in situations in which more than one book is present, that is, the property expressed by the adjective has to help to further specify the discourse referent, which is not possible if there is only one book. In the latter case it would be sufficient to use *il-ktieb* ('the book') to identify a certain book in a given context. This cardinality constraint seems to be correct for singular cases, but it is problematic for definite plural NPs, since they necessarily involve more than one discourse referent. The approach outlined by Kolliakou (2004) seems more promising. She suggests that adjectives that are marked as definite specify a subset of a contextually given set (Kolliakou, 2004, Section 5). She gives the specification in (25) for the polydefinite in (24):[6]

(24) to-kokino to-podilato
the-red the-bike
'the RED bike'

(25)
$$\begin{bmatrix} \text{CONTENT} & \begin{bmatrix} \text{INDEX} & \boxed{1} \\ \text{RESTR} & \{red(\boxed{1}), bike(\boxed{1})\} \\ & \textit{def-rel} \end{bmatrix} \\ \text{CONTEXT}|\text{BACKGROUND} & \left\{ \begin{bmatrix} \text{SUPERSET} & \boxed{2} \\ \text{SUBSET} & \boxed{1} \\ & \textit{include-in-rel} \end{bmatrix} \right\} \end{bmatrix}$$

Kolliakou uses Situation Semantics that was also used in (Pollard and Sag, 1987). The index $\boxed{1}$ in (25) refers to one element, not to a set. Hence the constraint $\boxed{1} \subset \boxed{2}$ is ill-formed. Instead one should say that the set containing $\boxed{1}$ has to be a subset of $\boxed{2}$. However, two important questions remain unanswered: what is $\boxed{2}$ and what happens in cases with more than one definite adjective. Kolliakou discusses the example:

[6] I replaced the set union sign in the restriction set by a comma. Kolliakou's description contains the constraint 'where $\boxed{1} \subset \boxed{2}$' in addition to what is given in (24), but this seems to be redundant, since this constraint is contained in the BACKGROUND set.

(26) to-kenurio to-kokino to-podilato
the.new the.red the.bike
'the new red bike'

I believe that a fusion of Fabri's and Kolliakou's analyses is the correct way to tackle this problem: instead of requiring that there is more than one entity that fits the predicate of the noun, one should assume a subset approach, that is, the definite adjective *il-ġdid* comes with the restriction that the set of possible discourse referents described by the noun is a proper superset of the set of discourse referents described by both the noun and the adjective:

(27) { book(x) } $\supset$ { book(x) $\wedge$ new(x) }

In situations in which there is just one book, the information that we are referring to a new book does not distinguish between possible discourse referents, and the use of the definite version of the adjective is therefore infelicitous. This subset approach also works for definite plural NPs.

A remaining open question is how to integrate the respective constraint into the grammar. Fabri suggests that the constraint should be attached to the empty determiner or to the demonstrative. The problem with this approach is that it invloves a certain nonlocality, since the determiner has to take information into account that can be deeply embedded in the nominal projection it is combined with. The alternative is to specify the constraint in the lexical item for the adjective. Kolliakou rejected this analysis since she assumes that the *to* of *to-kenurio* is the definite article in monadic definites like (28):

(28) to-kenurio podilato
the.new bike
'the new bike'

Instead she assumes a constructional analysis for polydefinites that introduces the subset condition on the phrasal level. Since I assume that the quantificational information is contributed by the demonstrative or by an empty element, Kolliakou's argument does not apply. However, due to the use of Minimal Recursion Semantics and due to the representation of the RELS list outside of SYNSEM, it is not possible to access the semantic contribution of a modified nominal projection from within the adjective.[7] Therefore locality considerations enforce a phrasal analysis, since the RELS values of daughters that are combined in a certain syntactic configuration can be accessed by the respective schema or general constraints on phrasal configurations.

Figure 3 on the following page shows an analysis that makes use of all the points discussed above. The adjective selects the noun via MOD. It can access the referential index of the noun (*x* in Figure 3) and identifies it with the index that is used in the RELS list of the adjective. The noun selects a determiner via SPR. The determiner is located in the position the demonstrative would take, but is not visible. The version of the determiner that is used in the analysis in Figure 3 selects a definite nominal projection via SPEC and contributes definite semantics.

[7] Copestake et al. (2005) assume a KEY feature that points to the main contribution of a head. This feature can be used to access the semantic contribution of the noun. However, in nominal phrases that involve multiple definite adjectives, this is not sufficient. According to Fabri (p.c. 2008) nominal structures with multiple definite adjectives are rather marked, so that one could decide that the grammar should not admit them at all. If this is the case, the KEY feature is sufficient and the subset condition can be stated lexically in the definite adjective.

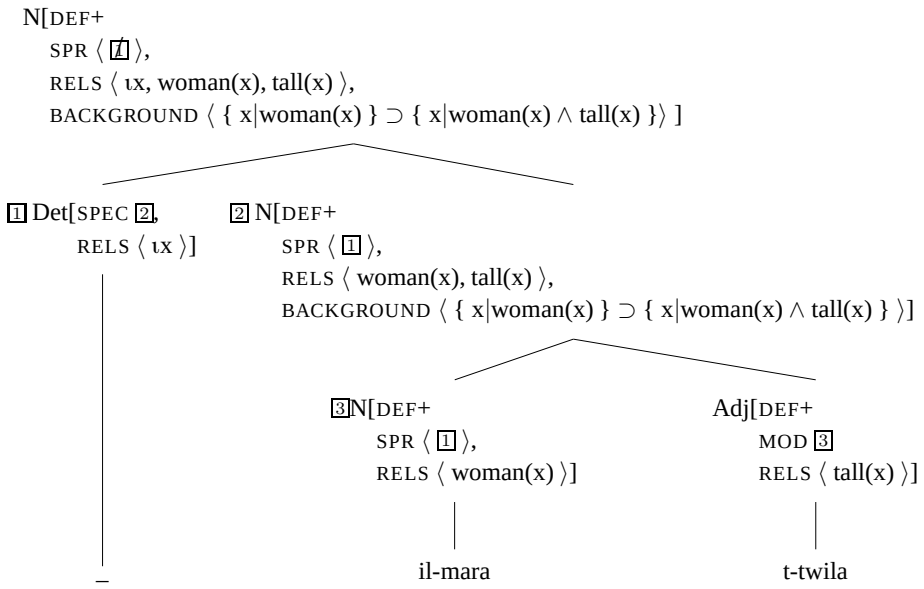


Figure 3: Analysis of *il-mara t-twila* 'def-woman def-tall' = 'the tall woman'

5 Predication

Adjectives can form a complete utterance together with their subject without involving a copula:

(29) a. Jien sabiħ.
I beautiful.msg
'I am beautiful.' (male speaker)

b. Jien sabiħ-a.
I beautiful-fsg
'I am beautiful.' (female speaker)

I first followed Fabri in assuming an empty copula that contributes tense and aspect information, but Utz Maas (p.c. 2007) has since convinced me that one should leave this information underspecified and assume that the hearer infers this information from the absence of any other marking. Therefore I represent the subject of predicative adjectives under SPR. The combination of *jien* and *sabiħa* is an instance of head-specifier phrase. For languages that need a copula, the subject of predicative adjectives is not presented under SPR, but under a head feature SUBJ. This SUBJ feature is not a valence feature and hence combinations like *Peter klug* (Peter smart) are ruled out. In order to form grammatical sentences with predicative elements, a copula has to be used that selects the predicative element and attracts the SUBJ value of the embedded predicate and turns it into a dependent of the copula. See Hinrichs and Nakazawa, 1989, 1994 for the general approach of argument attraction in HPSG and Müller, 2002, p. 103 for an analysis of the copula in German. A discussion of copulaless predication in German can be found in Müller, To Appear.

The copula can be omitted in the present tense only. For sentences with past tense a copula element has to be used:

(30) a. Pawlu tabib
Pawlu doctor
'Pawlu is a doctor.'

b. Pawlu kien tabib
Pawlu be.3msg.past doctor
'Pawlu was a doctor.'

The same is true for sentences with negation. These facts are captured by the analysis since it is assumed that the hearer infers the present tense from the absence of the copula. If a copula is present, nothing has to be inferred and the sentence is interpreted according to the form of the copula element. The negation needs something to attach to. This is the reason why the copula element is necessary in negated sentences.

One advantage of the copulaless analysis is that it is possible to restrict pro-drop (see Section 9) to verbs. Since no verb is involved in the analysis of (30), this explains why the subject cannot be dropped. If one would assume an empty verbal head for the analysis of (30), part of speech information could not be used to block pro-drop.

6 *Lil* Marking and Case Assignment

Subjects may not be *lil* marked. The direct object is *lil* marked if it refers to a human object (Fabri, 1993, Chapter 5.2):

(31) a. Raj-t lil Pawlu.
see-1sg Case Pawlu
'I saw Pawlu.'

b. * Raj-t Pawlu.
see-1sg Pawlu

c. Xtraj-t il-ktieb
buy-1sg def-book
'I bought the book.'

d. * Xtraj-t lil-l-ktieb
buy-1sg Case-def-book
'I bought the book.'

The indirect object is always *lil* marked:

(32) a. Bagħat il-ktieb lil Ġanni.
sent-3msg def-book Case Ġanni
'He sent the book to Ġanni.'

b. * Bagħat il-ktieb Ġanni.
sent-3msg def-book Ġanni

c. Bagħat l-ittr-a lil xi universita-jiet il-Ġermanja.
sent-3msg def-letter-fsg Case some university-pl def-Germany
'He sent the letter to some universities in Germany.'

A human direct object can be unmarked if a marked indirect object is present.

When proper names are *lil* marked, the proper name and *lil* are written separately (33), in all other cases the noun and *lil* are written together (34):

(33) lil Ġanni
Case Ġanni

(34) a. lil-l-mara
Case-def-woman

b. l-it-tifel
Case-def-boy

If *lil* attaches to an element that is marked for definiteness, *lil* is fused with the definiteness marker (34b).

Pronouns can appear either with or without *lil*. *lil* marked pronouns can be used as direct and indirect objects. Regardless of their status as direct or indirect object they always refer to animate entities. Therefore Fabri (1993, p. 123) suggests that these pronouns are not formed according to the general rules of *lil* marking, but are stored as fixed idiosyncratic items in the lexicon.

The combination of regular nouns with *lil* is taken care of by the following lexical rule:

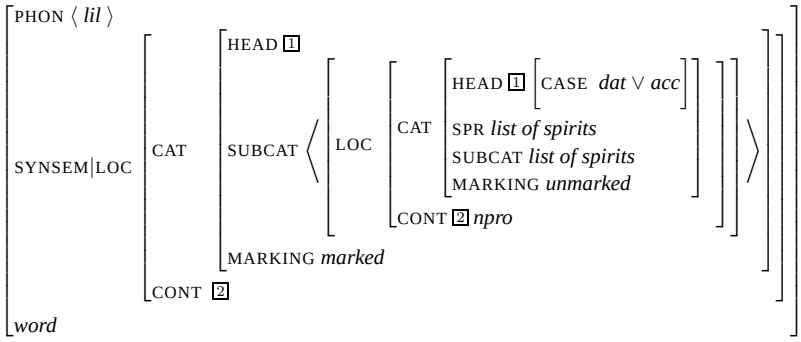
(35)
$$\left[\begin{array}{l}\text{PHON } \boxed{1} \\ \text{SYNSEM}|\text{LOC} \left[\begin{array}{ll}\text{CAT} & \left[\begin{array}{ll}\text{HEAD} & \left[\text{CASE } \textit{dat} \vee \textit{acc}\right] \\ \text{DEF} & \boxed{2} \\ \text{MARKING} & \textit{unmarked}\end{array}\right] \\ \text{CONT} & \textit{npro}\end{array}\right] \\ \text{RELS } \left\langle [\text{not } \textit{proper-noun-q}\ ] \right\rangle \oplus \square \\ \textit{word}\end{array}\right] \mapsto$$
$$\left[\begin{array}{l}\text{PHON f}(\boxed{1},\boxed{2}) \\ \text{SYNSEM}|\text{LOC}|\text{CAT}|\text{MARKING } \textit{marked}\end{array}\right]$$

The rule takes a noun as input that has the case value dat $\vee$ acc. This captures the fact that *lil* marked nouns cannot be used as subjects (nominative noun phrases). The rule furthermore requires that the semantic contribution of the noun is of type *npro* (non-pronominal). This ensures that the rule does not apply to pronouns (remember that pronouns are not compositional and therefore listed in the lexicon). Because of the spelling constraints that require that *lil* is separated by a blank from proper nouns, the rule may not be applied to proper nouns. This is ensured by the specification of the element in the RELS list. $\square$ stands for some arbitrary value. In (35) $\oplus \square$ indicates that there may be further elements in the RELS list of the input.

The computation of the phonological form has to take into account the definiteness marking of the input noun ($\boxed{2}$), since the *lil* is fused into the definiteness marker if it is present (Fabri, 1993, p. 114). The fact that the output of the lexical rule is marked is represented by the MARKING value *marked*.

lil marking of proper nouns is achieved by the following lexical entry for *lil*:[8]

[8]This lexical item has the same effect as markers in Pollard and Sag's analysis (1994). However, no head-marker schema is necessary, since the entry in (36) is combined with a noun by the head-argument schema.

(36)


lil selects an unmarked, non-pronominal NP in the dative or accusative. The CONT value is contributed by the head in head-argument structures. *lil* is the head but does not contribute semantically. This is represented by the fact that the CONT value of *lil* is identified with the CONT value of the selected NP (2).

When *lil* is combined with an NP, the Marking Principle (Pollard and Sag, 1994, S. 45) ensures that the MARKING value of *lil* is identical to the MARKING value of the whole phrase.

There are at least two ways to ensure the right distribution of *lil*: The first and most straight-forward one is to specify the MARKING values in the lexical entries. (37) gives an example for a specification of a ditransitive verb.

(37) $\langle \text{NP}[\textit{unmarked}], \text{NP}[(\textit{human} \wedge \textit{marked}) \vee \textit{unmarked}], \text{NP}[\textit{marked}] \rangle$

The alternative is less direct. Implicational constraints that refer to case values of arguments that are realized in certain syntactic configurations ensure that the arguments have the right marking and sortal values.

7 Clitics

Arguments can be realized by clitics. If both the direct object and the indirect object are realized as a clitic, the clitics appear in the order dO < iO. In addition to the clitic an argument can be realized as a full NP. If both the clitic and the full NP are realized in the sentence, the NP expresses a topic.

Clitics are attached to their host by a lexical rule, which marks the respective elements in the SUBCAT list of the verb, noun, or preposition as saturated.

8 Constituent Order

The subject can be placed to the left or to the right of the verb:

(38) a. Pawlu ġie.
Pawlu came

b. Ġie Pawlu.
came Pawlu

Fabri (1993, p. 138) shows that SV order is marked if the subject is indefinite and mentions various other factors that influence markedness of certain orders, but in principle both the SV and the VS order is attested in Maltese.

The position of the subject in transitive clauses is also rather free, if the subject is a topic. Without the object clitic there are the ordering variants in (39b, d–f) for the sentence in (39a) in reply to the utterance *U Ingrid?* (‘and Ingrid’):

(39) a. Ingrid kiel-et il-mazzit-a. SVO
Ingrid eat-3fsg def-black.pudding-fsg
‘Ingread ate black pudding.’

b. Kielet il-mazzita Ingrid. VOS

c. * Kielet Ingrid il-mazzita. VSO

d. Ingrid il-mazzita kielet. SOV

e. Il-mazzita Ingrid kielet. OSV

f. Il-mazzita kielet Ingrid. OVS

If the subject is focussed, it has to be placed at the left of the verb:

(40) Min feta ħ l-bieb?
who.msg open.3msg def-door.msg
‘Who opened the door?’

(41) a. Norma fet ħ-et il-bieb. SVO
Norma open-3fsg def-door.msg
‘Norma opened the door.’

b. * Fet ħet Norma l-bieb. VSO

c. * Norma il-bieb fet ħet. SOV

d. * Il-bieb fet ħet Norma. OVS

e. * Il-bieb Norma fet ħet. OSV

f. * Fet ħet il-bieb Norma. VOS

If the subject is focused and the object is realized both as a clitic and as a full NP, the following orders are possible (I use the small ‘o’ for clitics and the capital ‘O’ for full NPs):

(42) a. Norma fet ħ-it=u l-bieb. SVoO
Norma open-3fsg=3msg def-door.msg
‘Norma opened the door.’

b. Fet ħ-it=u Norma l-bieb. VoSO
open-3fsg=3msg Norma def-door.msg

c. Norma l-bieb fet ħ-it=u. SOVo
Norma def-door.msg open-3fsg=3msg

d. Il-bieb fet ħ-it=u Norma. OVoS
def-door.msg open-3fsg=3msg Norma

e. Il-bieb Norma fet ħ-it=u. OSVo
def-door.msg Norma open-3fsg=3msg

f. * Fet ħ-it=u l-bieb Norma. VoOS
open-3fsg=3msg def-door.msg Norma

With the subject topic and the object appearing as a clitic, we get:

(43) a. Ingrid kil-it=ha l-mazzita. SVoO
Ingrid ate-3fsg=3fsg def-black.pudding-fsg

b. Kil-it=ha l-mazzita Ingrid. VoOS
c. Kil-it=ha Ingrid il-mazzita. VoSO
d. Ingrid il-mazzita kil-it=ha. SOVo
e. Il-mazzita Ingrid kil-it=ha. OSVo
f. Il-mazzita kil-it=ha Ingrid. OVoS

Fabri (1993, p. 145) provides Table 1 as a condensed overview of the data:

focus subject			topic subject		
intransitive verb	transitive verb		intransitive verb	transitive verb	
semantic/ pragmatic restricted	without dO-clitic	with dO-clitic		without dO clitic	with dO clitic
	SVO only	*VoOS	absolutely free	*VSO	absolutely free

Table 1: Summary of possible constituent orders

As a reviewer pointed out to me the orders VOS and VoOS with a focused subject are also possible, albeit very marked. Thomas Stolz provided the following attested example of VOS:

(44) Hekk in-kun n-ista' n-aghzel xi bicca xoghol jien.[9]
thus 1sg-be.Fut 1sg-can 1sg-choose some piece work I
'I will be able to look for a job then.'

8.1 Subject Position

Following Fabri, I assume that Maltese is a configurational language, that is, a language that combines the verb with its complements to form a VP which is then combined with the subject to form a complete clause. As discussed in Section 2.2 this is modelled in HPSG by mapping complements and subjects to different lists: all complements are mapped to SUBCAT and the subject is mapped to SPR.

Since the head-specifier schema (see page 8) does not specify the order between subject (the element in the NON-HEAD-DTRS list) and VP (the head daughter) we get both orders, which is needed for intransitive verbs.

The order VSO without an object clitic is correctly excluded, since the subject can only combine with a VP, that is V and O have to be combined before the subject is combined with the result.

8.2 Object Fronting

Until now we can account for the patterns SV(O) and V(O)S. If one assumes that Maltese is a SVO language, other orders have to be related to this basic order. In what follows I will explain the analysis of OSV and SOV.

Building on work in the framework of GPSG (Gazdar, Klein, Pullum and Sag, 1985), Pollard and Sag (1994) developed an analysis for non-local dependencies. The

[9] Aldo FARINI (translated by Guzè CHETCUTI), Hrejjef ta' Gahan u stejjer ohra. Marsa: Dipartiment tat-Taghrif, 1997, p. 30.

basic idea is that a placeholder (trace) is used in the position in which a certain element is expected and that the information about the missing object is passed up in the tree until it is finally bound off at a higher node by its filler.[10] Figure 4 shows the details of the analysis of (39e), repeated here as (45) for convenience.

(45) Il-mazzita Ingrid kiel-et. OSV
def-black.pudding Ingrid ate-3fsg

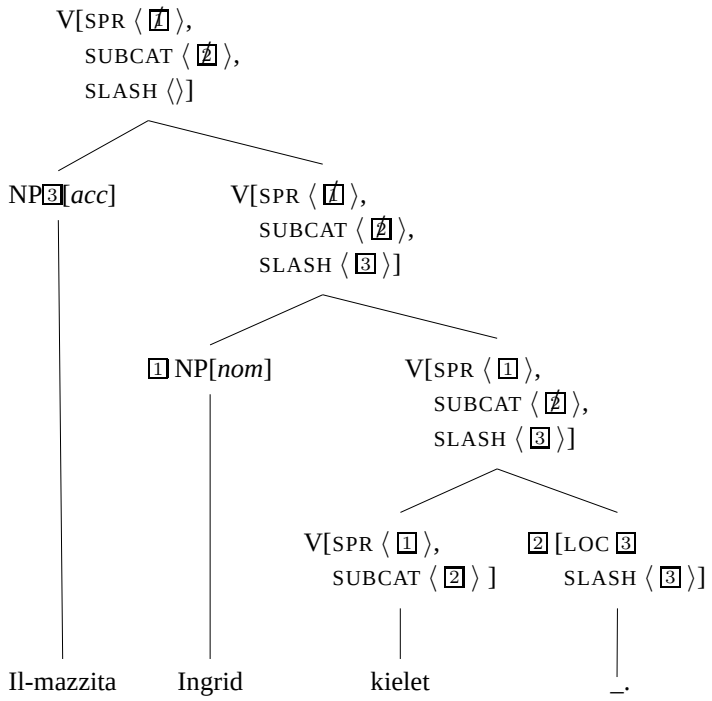


Figure 4: Analysis of the OSV order

The trace is combined with the verb to form a VP. The verb contains a description of the object that it requires in its SUBCAT list. This description is identified with the trace. Since the trace shares its local properties (those under LOC) with the element in SLASH, the information about the missing object is recorded. This information is passed up to the dominating nodes by the Nonlocal Feature Principle (Pollard and Sag, 1994, p. 164):

Principle 2 (Nonlocal Feature Principle)
For each nonlocal feature, the INHERITED *value of the mother is the union of the* INHERITED *values of the daughters minus the* TO-BIND *value on the head daughter.*

The top-most node in Figure 4 is licensed by the head-filler schema, that is given as Schema 4:

[10] There have been attempts to develop traceless analyses of nonlocal dependencies (Bouma, Malouf and Sag, 2001). The lexical variants have been criticized by Levine and Hukari (2006). Analyses that involve unary projections instead of empty elements are a notational variant of the analysis presented here. See also Müller, 2002, Chapter 6.2.5.1 and Müller, To Appear for discussion.

Schema 4 (Head-Filler Schema)
head-filler-phrase →

$$\begin{bmatrix} \text{HEAD-DTR}|\text{SYNSEM} & \begin{bmatrix} \text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & \textit{fin} \\ & \textit{verb} \end{bmatrix} \\ \text{SUBCAT} & \textit{list of spirits} \end{bmatrix} \\ \text{NONLOC}|\text{TO-BIND}|\text{SLASH} & \langle \boxed{1} \rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{NONLOC}|\text{INHER}|\text{SLASH} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

This schema combines a VP that is missing an element (the element in SLASH) with this missing element (the non-head daughter). Due to the nonlocal feature principle the information in SLASH is not passed to the mother node. The nonlocal dependency ends in the head filler phrase.

A linearization constraint ensures that the filler is serialized to the left of the non-head daughter.

Note that the schema differs from the one in Sag et al., 2003, p. 438 in not mentioning the SPR value of the head-daughter. This allows the head daughter to be a full clause or a VP. Therefore this schema accounts not only for the OSV order in (39e) but also for the SOV order in (39d).

8.3 Clitic Dislocation

As was discussed in the data section, full NPs can be used in addition to clitics on the verb. The phenomenon of clitic dislocation is also known from other languages. Alexopoulou and Kolliakou (2002) discuss Clitic Left Dislocation in Greek. The analysis adopted here is based on theirs. I assume that clitics introduce *local* objects under NONLOC|INHER|CLD. The information is projected to dominating nodes and can be bound off by full NPs. The schema that binds off elements in CLD is parallel to the head-filler schema:

Schema 5 (Head-CLD Schema)
head-cld-phrase →

$$\begin{bmatrix} \text{HEAD-DTR}|\text{SYNSEM} & \begin{bmatrix} \text{LOC}|\text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{VFORM} & \textit{fin} \\ & \textit{verb} \end{bmatrix} \\ \text{SUBCAT} & \langle\rangle \end{bmatrix} \\ \text{NONLOC}|\text{TO-BIND}|\text{CLD} & \langle \boxed{1} \rangle \end{bmatrix} \\ \text{NON-HEAD-DTRS} & \left\langle \begin{bmatrix} \text{SYNSEM} & \begin{bmatrix} \text{LOC} & \boxed{1} \\ \text{NONLOC}|\text{INHER} & \begin{bmatrix} \text{CLD} & \langle\rangle \\ \text{SLASH} & \langle\rangle \end{bmatrix} \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

In contrast to head-filler phrases, there is no restriction as far as the ordering of the daughters is concerned. The VP consists of the verb and a clitic. Given the right information structural context, the subject can be placed to the right and to the left of the VP, since the head-specifier schema allows for both orders. Therefore we can account for the orders in (46):[11]

[11] The last constituent order is restricted to subjects that are not focused.

(46) a. SOVo
b. OSVo
c. SVoO
d. OVoS
e. VoSO
f. VoOS (non-focussed subject)

8.4 Adjuncts

According to Fabri (1993, p. 138) adjuncts can be placed anywhere in the clause.

(47) a. Illum Pawlu ġie.
today Pawlu came
b. Pawlu ġie ’llum.
c. Pawlu illum ġie.
d. Ġie ’llum Pawlu.
e. Ġie Pawlu illum.

The following lexical entry for *illum* ‘today’ can account for this variability:

(48)
$$\begin{bmatrix} \textsc{phon} \langle\, \textit{illum} \,\rangle \\ \textsc{synsem|loc|cat} \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textsc{mod} & \textsc{v}\boxed{1} \\ \textit{adv} & \end{bmatrix} \\ \textsc{spr} & \langle\rangle \\ \textsc{subcat} & \langle\rangle \end{bmatrix} \\ \textsc{rels} \left\langle \begin{bmatrix} \textsc{arg1} & \boxed{1} \\ \textit{today} & \end{bmatrix} \right\rangle \end{bmatrix}$$

Since neither the valence properties nor any nonlocal properties (SLASH or CLD) of the modifiee are specified, the adverb can attach to all verbal projections. Since the PRE-MODIFIER value of the adverb is not specified (compare the lexical entry for *sabiħ* in (21). This adjective obligatorily follows the noun it modifies), it can be placed to the left and to the right of the verbal projection it modifies.

8.5 A Technical Problem: Spurious Ambiguities

There is a technical problem that is caused by the fact that the clitic dislocation schema has to allow both SOVo and OSVo. Because of the SOVo order, the schema cannot require that the dislocated NP attaches to a fully saturated verbal projection. While the flexibility is desired if both the dislocated NP and the subject are located to the same side of the head we get spurious ambiguities when they are located on different sides. Figure 5 on the following page shows an example of such a situation.

Spurious ambiguities of this type can result in a combinatorial explosion if adjuncts and the pro drop rule (see Section 9) are involved. (49) shows an example involving adverbs. Since S Vo Adv O is possible, (49b) cannot be ruled out in general by requiring that adverbs attach only if all dislocated clitics are bound off.

(49) a. [O S Vo] Adv

Figure 5: The problem of spurious ambiguities

b. O [S Vo Adv]

The solution is to use a feature that marks a projection if some constituent is right-adjoined to it. Schemata that left-adjoin material require that the head-daughter does not contain any right-adjoined material yet. So we get the left structure in Figure 5 and rule out the right structure. The left structure is the one that is cognitively the more plausible one, since humans start to build structure as soon as they hear material.

9 Pro-Drop

In contrast to the object clitics, the subject properties are obligatorily marked on the verb. Clitics fill the argument slot of a verb. Therefore the arguments do not have to be realized as a full NP. Since the marking of the verb with respect to the person and number features is obligatory, I treat this marking as inflection. However, the realization of the subject is optional. In principle, there are several possibilities to account for this optionality: an empty pronoun that appears in the position in which an overt subject would be realized, a lexical rule that reduces the valence of the verb, a unary branching rule that discharges the specifier element of a VP. I do not see any empirical differences between the three approaches and have therefore decided to choose the one with the unary branching rule, since it has the best properties in terms of the computer implementation. The schema is given in (50):

Schema 6 (Pro-Drop Schema)
pro-drop-phrase →

$$\left[\begin{array}{ll} \textsc{synsem}|\textsc{loc}|\textsc{cat}|\textsc{spr} \langle \boxed{1} \rangle & \\ \textsc{head-dtr}|\textsc{synsem}|\textsc{loc}|\textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \textit{verb} \\ \textsc{spr} & \langle \boxed{1}\, \mathrm{NP}_{ppro} \rangle \\ \textsc{subcat} & \textit{list of spirits} \end{array}\right] \end{array}\right]$$

The head daughter has to be a VP, that is an object of the category verb that has only spirits in its SUBCAT list and one element in its SPR list. The schema instantiates this element to be a personal pronoun. Due to subject verb agreement, we know the person and number values of the dropped subject and hence the reference can be reconstructed by taking discourse information into account.

10 Summary and Outlook

This paper describes an implemented fragment of Maltese grammar. The fragment can be downloaded at http://hpsg.fu-berlin.de/Fragments/Maltese/. The grammar fragment shares a common core with implemented grammars of German, Persian, and Mandarin

Chinese. Important parts of the Maltese language are covered, but some crucial bits are still missing: The parts of morphology that involve vowel changes are not yet implemented. While this is neither a theoretical nor an implementational problem for the framework and system used, it involves setting up the vowel patterns and classification of stems, work that should be done by a native speaker. The grammar is lacking rules for active/passive alternations, causative constructions, complex predicates, raising and control, and numerals in the NP. Many of these phenomena are already described in Fabri's work and should be implementable without problems.

In a more recent study, Fabri and Borg (2002) examined the constituent order data in more detail. This study has not been taken into account in the present implementation. It could turn out that a non-configurational treatment of Maltese is more appropriate. This would be easier to model than the fixed constituent order + dislocation. Instead of a head-argument schema that allows only the combination of a head with its most oblique non-realized argument, a more general schema could be used that combines any unrealized argument with the head. This is the treatment of constituent order that I assume for German (Müller, 2005, 2008). Of course constituent order interacts with information structure. There is promising work on information structure and the interfaces to phonology, syntax, and semantics in HPSG (Engdahl and Vallduví, 1994; De Kuthy, 2002; Bildhauer, 2008) and this research is being continued in connection with the Berlin/Potsdam Sonderforschungsbereich on information structure. Therefore, the research done on Maltese constituent order can be combined with work on constituent order and information structure being done in other projects.

References

Alexopoulou, Theodora and Kolliakou, Dimitra. 2002. On Linkhood, Topicalization and Clitic Left Dislocation. *Journal of Linguistics* 38(2), 193–245.

Bildhauer, Felix. 2008. *Representing Information Structure in an HPSG Grammar of Spanish*. Ph. D.thesis, Universität Bremen.

Borsley, Robert D. 1999. *Syntactic Theory: A Unified Approach*. London: Edward Arnold, second edition.

Bouma, Gosse, Malouf, Robert and Sag, Ivan A. 2001. Satisfying Constraints on Extraction and Adjunction. *Natural Language and Linguistic Theory* 19(1), 1–65. ftp://csli-ftp.stanford.edu/linguistics/sag/bms-nllt.ps, 18.08.2002.

Copestake, Ann, Flickinger, Daniel P., Pollard, Carl J. and Sag, Ivan A. 2005. Minimal Recursion Semantics: an Introduction. *Research on Language and Computation* 4(3), 281–332. http://lingo.stanford.edu/sag/papers/copestake.pdf, 11.10.2006.

De Kuthy, Kordula. 2002. *Discontinuous NPs in German*. Stanford, CA: CSLI Publications.

Engdahl, Elisabet and Vallduví, Enric. 1994. Information Packaging and Grammar Architecture: A Constraint-Based Approach. In Elisabet Engdahl (ed.), *Integrating information structure into constraint-based and categorial approaches*, pages 39–79, Amsterdam: ILLC, DYANA-2 Report R.1.3.B.

Fabri, Ray. 1993. *Kongruenz und die Grammatik des Maltesischen*. Linguistische Arbeiten, No. 292, Tübingen: Max Niemeyer Verlag.

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

Fabri, Ray and Borg, Albert. 2002. Topic, Focus and Word Order in Maltese. In Abderrahim Youssi, Fouzia Benjelloun, Mohamed Dahbi and Zakia Iraqui-Sinaceur (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.

Gazdar, Gerald, Klein, Ewan, Pullum, Geoffrey K. and Sag, Ivan A. 1985. *Generalized Phrase Structure Grammar*. Cambridge, Massachusetts: Harvard University Press.

Higginbotham, James. 1985. On Semantics. *Linguistic Inquiry* 16(4), 547–593.

Hinrichs, Erhard W. and Nakazawa, Tsuneko. 1989. Subcategorization and VP Structure in German. In *Aspects of German VP Structure*, SfS-Report-01-93, Eberhard-Karls-Universität Tübingen.

Hinrichs, Erhard W. and Nakazawa, Tsuneko. 1994. Linearizing AUXs in German Verbal Complexes. In John Nerbonne, Klaus Netter and Carl J. Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, CSLI Lecture Notes, No. 46, pages 11–38, Stanford, CA: CSLI Publications.

Hudson, Richard. 2004. Are determiners heads? *Functions of Language* 11(1), 7–42. http://www.phon.ucl.ac.uk/home/dick/dets.htm, 29.05.2005.

Kiss, Tibor. 1992. Variable Subkategorisierung. Eine Theorie unpersönlicher Einbettungen im Deutschen. *Linguistische Berichte* 140, 256–293.

Kiss, Tibor. 1995. *Infinite Komplementation. Neue Studien zum deutschen Verbum infinitum*. Linguistische Arbeiten, No. 333, Tübingen: Max Niemeyer Verlag.

Kolliakou, Dimitra. 2004. Monadic Definites and Polydefinites: Their Form, Meaning and Use. *Journal of Linguistics* 40(2), 263–323.

Levine, Robert D. and Hukari, Thomas E. 2006. *The Unity of Unbounded Dependency Constructions*. CSLI Lecture Notes, No. 166, Stanford University: CSLI Publications.

Melnik, Nurit. 2006. Hybrid Agreement as a Conflict Resolution Strategy. In Stefan Müller (ed.), *The Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar*, pages 228–246, Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/7/, 05.20.07.

Meurers, Walt Detmar. 1999. Raising Spirits (and Assigning Them Case). *Groninger Arbeiten zur Germanistischen Linguistik (GAGL)* 43, 173–226. http://www.sfs.uni-tuebingen.de/~dm/papers/gagl99.html, 18.04.2000.

Meurers, Walt Detmar. 2000. Lexical Generalizations in the Syntax of German Non-Finite Constructions. Arbeitspapiere des SFB 340 No. 145, Eberhard-Karls-Universität, Tübingen. http://www.ling.ohio-state.edu/~dm/papers/diss.html, 19.08.2002.

Meurers, Walt Detmar, Penn, Gerald and Richter, Frank. 2002. A Web-Based Instructional Platform for Constraint-Based Grammar Formalisms and Parsing. In Dragomir Radev and Chris Brew (eds.), *Effective Tools and Methodologies for Teaching NLP and CL*, pages 18–25, proceedings of the Workshop held at 40th Annual Meeting of the Association for Computational Linguistics. Philadelphia, PA. http://www.ling.ohio-state.edu/~dm/papers/acl02.html, 08.01.2004.

Müller, Stefan. 1999. *Deutsche Syntax deklarativ. Head-Driven Phrase Structure Grammar für das Deutsche*. Linguistische Arbeiten, No. 394, Tübingen: Max Niemeyer Verlag. http://hpsg.fu-berlin.de/~stefan/Pub/hpsg.html, 09.03.2010.

Müller, Stefan. 2002. *Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German*. Studies in Constraint-Based Lexicalism, No. 13, Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/complex.html, 09.03.2010.

Müller, Stefan. 2005. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203, 297–330. http://hpsg.fu-berlin.de/~stefan/Pub/mehr-vf-lb.html, 09.03.2010.

Müller, Stefan. 2008. *Head-Driven Phrase Structure Grammar: Eine Einführung*. Stauffenburg Einführungen, No. 17, Tübingen: Stauffenburg Verlag, second edition. http://hpsg.fu-berlin.de/~stefan/Pub/hpsg-lehrbuch.html, 09.03.2010.

Müller, Stefan. To Appear. Elliptical Constructions, Multiple Frontings, and Surface-Based Syntax. In Gerhard Jäger, Paola Monachesi, Gerald Penn and Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004, Nancy*, Stanford, CA: CSLI Publications. http://hpsg.fu-berlin.de/~stefan/Pub/surface.html, 09.03.2010.

Müller, Stefan and Vasishth, Shravan. In Preparation. *Hindi Syntax in Head-Driven Phrase Structure Grammar*.

Penn, Gerald. 2004. Balancing Clarity and Efficiency in Typed Feature Logic Through Delaying. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), Main Volume*, pages 239–246, Barcelona, Spain.

Pollard, Carl J. and Sag, Ivan A. 1987. *Information-Based Syntax and Semantics*. CSLI Lecture Notes, No. 13, Stanford, CA: CSLI Publications.

Pollard, Carl J. and Sag, Ivan A. 1994. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics, Chicago, London: University of Chicago Press.

Przepiórkowski, Adam. 1999a. *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph. D.thesis, Eberhard-Karls-Universität Tübingen, Germany. http://w210.ub.uni-tuebingen.de/dbt/volltexte/1999/72/, 19.11.2005.

Przepiórkowski, Adam. 1999b. On Case Assignment and "Adjuncts as Complements". In Gert Webelhuth, Jean-Pierre Koenig and Andreas Kathol (eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, Studies in Constraint-Based Lexicalism, No. 1, pages 231–245, Stanford, CA: CSLI Publications.

Sag, Ivan A., Wasow, Thomas and Bender, Emily M. 2003. *Syntactic Theory: A Formal Introduction*. CSLI Lecture Notes, No. 152, Stanford, CA: CSLI Publications, second edition.

Sailer, Manfred. 2004. Local Semantics in Head-Driven Phrase Structure Grammar. In Olivier Bonami and Patricia Cabredo Hofherr (eds.), *Empirical Issues in Formal Syntax and Semantics*, volume 5, pages 197–214, Online. http://www.cssp.cnrs.fr/eiss5/sailer/index_en.html, 23.09.2007.

van Langendonck, Willy. 1994. Determiners as Heads? *Cognitive Linguistics* 5, 243–259, vanEynde: gegen DP, Hudson2004: gegen DP.

Wechsler, Stephen Mark and Zlatić, Larisa. 2003. *The Many Faces of Agreement*. Stanford Monographs in Linguistics, Stanford, CA: CSLI Publications.

Wunderlich, Dieter. 1997. Cause and the Structure of Verbs. *Linguistic Inquiry* 28, 27–68.

Yip, Moira, Maling, Joan and Jackendoff, Ray S. 1987. Case in Tiers. *Language* 63(2), 217–250.