1. Introduction

I chose the Italian verb morphology because of personal interest. I studied the language in school but did not have, or rather did not take the opportunity to speak it much the last few years. Therefore, I thought it might be a good idea to use an Italian phenomenon to find my way back into the language. During the research and implementation, I realized that not all was lost and I started to remember learning the conjugations of verbs in school.

Italian is a romance language and is derived from Latin. It developed over time starting with vulgar Latin, also called colloquial Latin. Due to the size of the Roman empire many different cultures and languages influenced colloquial Latin and the different romance languages were created. One of them is, after a long history that will not be the topic here, Italian as it is spoken today (Reutner & Schwarze, 2011). Even though it is its own language today, its origin in Latin can still be seen clearly. This also reflects in the rich morphology which is one of the characteristics of Latin. Many forms have been eliminated over time during the development of colloquial Latin and suffixes became indistinguishable in phonetic evolution resulting in the slightly less rich morphology in Italian (Jensen, 1971). But the combination of a language with a great variety in morphology and it being spoken today makes it a great project for this course.

I wrote an Italian grammar from scratch and decided to include a number of morphosyntactic phenomena, which I'll explain in the following.

2. Implemented phenomena

Agreement:

As mentioned earlier, morphology is far richer in Italian than it is in English. That means we have to include gender, number, person and mood information in more detail. Especially adjectives and determiners have to agree with the nouns they modify (e.g., *la nave nera* = *the black ship* vs. *il libro nero* = *the black book*). This has to be accounted for in our grammar. I solved this issue via a combination of information in the lexicon and rules in the c-structure. In the lexicon I included gender and number information via a template for adjectives that are variable when it comes to gender (see Figure 1). There are other ones that only are variable in number, which I represent via another template without gender information (see Figure 2).

```
ADJECTIVE-AGR (P G N D T) = (^ PRED) = 'P'
                            (^ GEN) = G
                            (^ NUM) = N
                            (^ DEGREE) = D
                            (^ APOS) = T
                            (^ ADJ-TYPE) = adjective.
```

Figure 1: Template of adjective with agreement

```
nera            A * @(ADJECTIVE-AGR nero fem sg positive post).
```

Figure 2: Sample entry of adjective with agreement (nera = black)

But having this information in the lexicon alone is not enough. It needs to be passed up to the NP constituent, so we can guarantee feature unification and make sure that nouns and adjectives that don't agree get an ungrammatical judgement.

```
{AP*: ! $ (^ ADJUNCT)
          (^ GEN) = (! GEN)
          (^ NUM) = (! NUM)
          (! APOS) =c pre;
```

Figure 3: Prenominal adjective phrase in the NP rule

Working and ungrammatical examples for this can be seen in test suite sentences 1 – 12.

Adjective placement:

Something else that can be seen in the rules and the lexical entries is the fact that in Italian adjectives can occupy different positions. Lexical ones like colours are in postnominal position, while functional ones like first, second, etc. are in prenominal position (Ramaglia, 2008). To account for this in our grammar, we again use a combination of lexical information (*pre* and *post* as a feature for each adjective) and constraints in the syntactical rules. For an adjective to be in prenominal position it has to have the value *pre* for *APOS* as can be seen in Figure 3. There is one additional case, where some adjectives can appear both before and after the verb, resulting in a slight meaning change, which will not be discussed here as it goes beyond the scope of this project (for a discussion see Ramaglia, 2008). That's why there is a disjunction in the NP rule for {AP N} or {N AP}. Adjectives that can occupy both positions then have a disjunction themselves in their lexical entries, so they can appear in both before and after the noun (see Figure 4).

```
vecchio          A * {@(ADJECTIVE-AGR vecchio masc sg positive pre)
                     |@(ADJECTIVE-AGR vecchio masc sg positive post)}.
```

Figure 4: Sample entry of an adjective that can occupy both positions around the noun

The disjunction in the c-structure ensures that these adjectives can only be present in one of the two positions and not in both at the same time (see test sentences 13 – 19)

Adverbs:

Adverbs work slightly differently in Italian as they appear in postverbal position in Italian, when modifying a verb, compared to English, where they can appear both before and after the verb. Unlike Italian adjectives, adverbs are invariable, meaning they don't carry any gender, number or person information. That is why the template for this word class is rather simple (see Figure 5).

```
ADVERB(P D)      =    (^PRED) = 'P'
                      (^DEGREE) = D
                      (^ADJ-TYPE) = adverb.
```

Figure 5: Template for adverbs

In the c-structure rules I specify the fact that they can only appear postverbal in the VP rule (see Figure 6).

```
(AUX)
V: ^=!;
ADVP*: ! $ (^ ADJUNCT);
```

Figure 6: Postverbal adverb phrase in VP rule

When modifying adjectives, they are more flexible and are accounted for in the AP rule (see Figure 7).

```
AP  -->      ADV*: ! $ (^ ADJUNCT);
             A
             (COMMA).
```

Figure 7: Adjective phrase rule

Examples with adverbs can be seen in 20 – 25 in the testsuite.

Passive:

Another feature that is included in the grammar, is passive sentences. This is implemented via a template like in our English grammar (Figure 8). For it to apply, we need a verb in a past participle form, which is specified in the lexicon for the corresponding verb forms. If this is given, we pass the

information that we have a passive sentence up and change the object to the subject and the subject to an oblique agent or delete it.

```
{ FRAME
| (^ PASSIVE) = -
| FRAME
  (^ PASSIVE) = +
  (^ PARTICIPLE) =c past
  (^ OBJ)-->(^ SUBJ)
{ (^ SUBJ)-->(^ OBL-AG)
| (^ SUBJ)--> NULL
}
}.
```

Figure 8: Template for passive sentences

In cases where we don't delete the subject, we need a prepositional phrase for the oblique agent. In Italian this is introduced by the preposition *da*. We specify such a constituent in the c-structure rules and constrain it to be of PFORM *da*. This then resembles our oblique agent constituent (see Figure 9).

```
(PP: (^ OBL-AG) = !
     | (! PFORM) =c da )
```

Figure 9: Passive PP in the VP rule

Examples of passive sentences can be found in the testsuite sentences 26 – 30.

Prepositional phrases:

Also similar to our English grammar, I have implemented prepositional phrases in the Italian grammar. There are two types, namely semantic ones and non-semantic ones. Semantic ones are usually found in adjuncts that add additional information to the sentence outside of the verb arguments. These can be found in both the VP and NP rules and are constraint via a constraining equation to have semantic prepositions (see Figure 12).

```
PP*: { ! $ (^ ADJUNCT)
         (! PTYPE) =c sem
         @(OT-MARK PPadj)
```

Figure 10: Prepositional phrase of semantic type

In the template for prepositions, we implement the semantic use as the preposition being the predicate and what follows it as its argument. We additionally add two OT marks that prefer these semantic PPs to be combined with not the VP (PPadj) but the NP (+PPattach), resulting in a modification of the latter.

The non-semantic PP can serve as an argument of the verb and loses its semantic value. This for example happens in passive sentences, where the oblique agent represented by a PP introduced by *da* constitutes a verb argument.

Examples for this can be found in 31 – 34 in the testsuite.

Ditransitive verbs:

I included ditransitive verbs in the grammar similar to the ones in English. In Italian these work a bit different though, since the indirect object is represented with a prepositional phrase introduced by *a*. Therefore, we need a different rule in our VP, where this argument is a PP with the P-FORM *a*. I constrain this in the c-structure as can be seen in Figure 13.

```
(PP: (^ OBJ-2) = !
     (! PFORM) =c a)
```

Figure 11: Prepositional phrase for ditransitives

Examples for the ditransitive verb dare (to give) can be found in the test sentences 35 – 38.

Pro-drop:

A phenomenon that can be seen in a few romance languages is pronoun dropping, where a pronoun can be left out, since the verb form itself already contains the necessary person and number information (Dalrymple, 2023). In order to implement this in our grammar we need to do two things. First, we have to create a c-structure rule that allows for an empty subject position. This is done similarly to the imperative structure, which can also be seen in sentences 39 – 45 in the testsuite. It constitutes its own sentence category (see Figure 11). Second, we need to incorporate the pronoun information in our verb entries. We will need one pure agreement form that takes any subject and object as arguments, and one pro-drop form, where the verb entry itself contains the pronoun information, making the actual pronoun superfluous (see Figure 10). In the latter case we pass a pro form up as the subject predicate within the lexical entry so that all argument positions of the verb are still filled, but we can still parse sentences with dropped pronouns (Dalrymple, 2023).

```
apre        V * {@(TRANS aprire)
                @VPRES
                @(VERB-AGR 3 sg)
                |
                @(TRANS aprire)
                @VPRES
                @(VERB-AGR 3 sg)
                (^ SUBJ PRED) = 'pro'}.
```

Figure 12: Lexical entry with pronoun incorporation

```
Sdrop -->   e
            VP: ^=!;
            (PERIOD).
```

Figure 13: Sdrop root category rule

Possessive pronoun constructions:

I also included possessive pronoun constructions (e.g., *Franca eats my pear.* = ***Franca mangia la mia mela.***). In comparison to English, Italian requires there to be determiner before the possessive pronoun. This has to be implemented in the c-structure and the agreement rules discussed earlier apply as well. Additionally, we need to specify that only possessive pronouns can appear in this position, which is ensured via two constraining equations (see Figure 14). These provide that only genitive possessive pronouns can be realised here.

```
| D
  PRON: (^ PRON)= !
        (! CASE) =c gen
        (! PRON-TYPE) =c poss
```

Figure 14: c-structure for possessive pronoun constructions

Examples of this morpho-syntactic construction can be found in 46 – 50 in the testsuite.

Tenses:

Lastly, I added more tenses. Overall, the parser can now parse present, imperfect and present perfect tenses. These can be found in the f-structure under the TENSE attribute, specified in the verb. A more complicated case is given in the present perfect because it is a combination of an auxiliary and a full verb. The auxiliary here carries the number and person information, while the full verb is invariable (at least in the cases with *avere* as the auxiliary). The verb *avere* can be a verb on its own though, which is why it has two entries in the lexicon. One with a predicate value (V) and one without (AUX), as can be seen in Figure 15.

```
ha  V   * @(TRANS avere)
          @VPRES
          @(VERB-AGR 3 sg);

    AUX * @VPRESPER
          @(VERB-AGR 3 sg).
```

Figure 15: One form of the verb avere with both V and AUX function

One problem that arises here is that the past participle and present perfect form of full verbs are identical. To prevent ungrammatical parses, where a passive structure is parsed as an active one, we add an OT-marker to prefer the present perfect reading of these forms in active sentences (see Figure 16).

```
scritto     V * @(TRANS scrivere)
                {@(PASTP masc)
                @(VERB-AGR 3 sg)
                |@VPRESPER
                @(OT-MARK presperfect)}.
```

Figure 16: OT-marker in present perfect verb forms

This ensures that both use cases work as can be seen in the test sentences 51 – 56 in the testsuite.

3. Files

I did not use a morphological analyser nor any other file, since I could not find any resources for Italian. Therefore, the lexicon is a full form lexicon. I did use an xlerc file to enable quick parsing of the testsuite and having the parser ready instantly. The test sentences are included in an extra file.

4. Future challenges

There is plenty of morpho-syntactic phenomena that could still be implemented. One thing that is very typical in Italian, but is not yet functional in the current grammar is the contraction of prepositions and definite articles. In our example xy in la would actually be nella. This could be something that can be implemented in future stages of the parser. Something else that can be accounted for later on, is the semantic difference depending on adjective placement discussed earlier. This could potentially be solved via the lexical entries and adding the meaning detail there.

5. Literature

Butt, M. (1999). *A grammar writer's cookbook*. CSLI Publications.

Dalrymple, M. (2023). *Handbook of Lexical Functional Grammar*. Language Science Press.

Jensen, F. (1971). *The Italian Verb*. The University of North Carolina Press.

Ramaglia, F. (2008). *La sintassi degli aggettivi e la proiezione estesa del NP*. Università degli Studi RomaTre.

Reutner, U., & Schwarze, S. (2011). *Geschichte der italienischen Sprache: Eine Einführung*. Narr.