Standardisation and Interoperation of Morphosyntactic and Syntactic Annotation Tools for Spanish and their Annotations

Antonio Pareja-Lora[*,**]**, Guillermo Cárcamo-Escorza**[*]**, Alicia Ballesteros-Calvo**[*]

[*] Universidad Complutense de Madrid (UCM) ; [**] ATLAS (UNED)

Ciudad Universitaria. 28040 – Madrid.

E-mail: aplora@ucm.es, guilleguiden@gmail.com, alibal01@ucm.es

Abstract

Linguistic annotation tools and linguistic annotations are scarcely syntactically and/or semantically interoperable. Their low interoperability usually results from the number of factors taken into account in their development and design. These include (i) the type of phenomena annotated (either morphosyntactic, syntactic, semantic, etc.); (ii) how these phenomena are annotated (e.g., the particular guidelines and/or schema used to encode the annotations); and (iii) the languages (Java, C++, etc.) and technologies (as standalone programs, as APIs, as web services, etc.) used to develop them.

This low level of interoperability makes it difficult to reuse both the linguistic annotation tools and their annotations in new scenarios, e.g., in natural language processing (NLP) pipelines. In spite of this, developing new linguistic tools from scratch is quite a high time-consuming task that also entails a very high cost.

Therefore, cost-effective ways to systematically reuse linguistic tools and annotations must be found urgently. A traditional way to overcome reuse and/or interoperability problems is standardisation. In this paper, we present a web service version of FreeLing that provides standard-compliant morpho-syntactic and syntactic annotations for Spanish, according to several ISO linguistic annotation standards and standard drafts.

Keywords: POS tagging, syntactic annotation, interoperability, standardisation, ISO, MAF, SynAF, web service, FreeLing.

1. Introduction

There are many conflicts and problems that prevent linguistic annotation tools and annotations from interoperating. Mostly, these conflicts and problems come from the number of factors that are taken into account in their development and design. As shown in Pareja-Lora (2012), linguistic annotation tools usually differ from each other either in (i) the type of phenomena they annotate – for instance, POS taggers annotate morphosyntactic phenomena, such as token segmentation or token features (like case or gender), whereas parsers provide some syntactic annotations, such as the constitution and/or dependency relationships that exist between the tokens of a text; (ii) how they annotate these phenomena, that is, the particular guidelines, meta-model, schema, formal language and tagset used to encode the annotations; and (iii) the languages (Java, C++, etc.) and technologies used to develop them (as standalone programs, as APIs, as web services, etc.).

These differences in the criteria followed to develop linguistic annotation tools often make it very difficult for them and their annotations to interoperate and be easily reused in new scenarios, e.g., in natural language processing (NLP) pipelines (Buyko et al., 2008). In spite of this, developing new linguistic tools from scratch is quite a high time-consuming task that also entails a very high cost. Therefore, the need to reuse the existing linguistic tools and find cost-effective ways to make them and/or their annotations interoperate gets clearer and more urgent every day.

A traditional way to overcome this reuse and/or interoperability problem in several areas (for instance, the (inter)connection of electronic plugs and/or devices) is standardisation. This is one of the main assumptions

driving the development of ISO linguistic annotation standards, such as ISO/MAF (2012) or ISO/SynAF (2010) and standard drafts, such as ISO/SynAF-<tiger2/> [1] (Bosch et al., 2012).

This is also the main hypothesis underlying the work presented in this paper: that standardisation can help linguistic tools and annotations interoperate (Ballesteros-Calvo et al., 2013). In particular, we wanted to test whether the standardisation of (1) a linguistic annotation tool (i.e., FreeLing) and (2) its morphosyntactic and syntactic annotations for Spanish, could help link these two types of annotations together and make them interoperate.

In the coming sections, we present how we accomplished this twofold standardisation task and the results obtained as for the interoperation of the standardised annotations. Thus, the rest of this paper has been organised as follows. Firstly, we introduce FreeLing and the reasons that lead us to standardise this linguistic annotation tool and its outputs. Secondly, we discuss how this standardisation was accomplished, focusing on the standardisation of (1) its morphosyntactic annotations and (2) its syntactic annotations. Thirdly, we state the conclusions of this research. Finally, we have included the acknowledgements and the references associated to this work.

2. Why FreeLing

FreeLing (Padró & Stanilovsky, 2012) is an open source, freely available tool for the analysis and annotation of texts at several levels and layers, written in a number of languages (see below). The version standardized in this

---

[1] An XML serialization of ISO/SynAF (2010) that is based on the TIGER (König & Lezius, 2003) language and format.

work (FreeLing 3.0) processes the input text and provides, for instance, its (i) token segmentation, (ii) POS tagging, (iii) deep and shallow syntactic constituency parsing; (iv) syntactic dependency parsing, (v) multiword detection, (vi) named entity recognition and classification (according to the MUC classification – Chinchor, 1997), and (vi) (Euro)WordNet-based (Miller, 1995; Fellbaum, 1998; Vossen, 1998) sense tagging.

The input text can be written in Spanish, Catalan, Galician, Portuguese, Italian, French, English or Russian, amongst others. Its openness, availability, versatility and multilingual capabilities have made it a very popular and most widespread tool, for instance, in Spain.

However, FreeLing has a few limitations that require being solved. For example, on the one hand, none of its manifold output annotations comply with the current standards for linguistic annotation (such as ISO/MAF (2012) or ISO/SynAF (2010)) and are not even encoded by means of a standard language (such as XML[2]). On the other hand, its current implementation does not allow for its inclusion 'as is' into NLP pipelines. These two factors altogether reduce to some extent the interoperability and reusability of this tool.

3. Standardizing FreeLing

So, in order to overcome all these problems, first of all, we decided to transform FreeLing 3.0 into a web service[3], which is a fairly well-known, widespread and standard-based[4] way to improve the interoperability of computer applications (Kashyap et al., 2008)[5].

In spite of this first standardisation step, FreeLing's results (i.e., its annotations) were still encoded in a tool-dependent, non-standard-compliant way. For example, the tokens in its POS annotations were not assigned a URI in order to allow other annotations of the same text to refer to them. This made it difficult to (i) interconnect its annotations together and (ii) merge them with the annotations performed by other tools. This, in turn, prevented the tool from being sufficiently interoperable. Accordingly, a second step towards the standardisation of FreeLing was required, namely the standardisation of its annotations.

Previous approaches[6] (Poch & Bel, 2011; Morell, Vivaldi & Bel, 2012) had already accomplished the standardisation of FreeLing's (morpho-)syntactic outputs using the Graph Annotation Format (GrAF: Ide & Suderman, 2007). GrAF is an XML serialization of the standard Linguistic Annotation Framework of ISO (ISO/LAF, 2012). ISO/LAF (2012) and GrAF altogether provide a general annotation framework, pretty suitable

for those cases for which no other ISO annotation standard is available. However, it is too general for quite common and useful types of annotations, such as morpho-syntactic and syntactic annotations, for which other specific ISO annotation standards have already been developed (namely ISO/MAF (2012) and ISO/SynAF (2010), respectively). While these other ISO standards are also ISO/LAF-compliant, (a) they are also less verbose than GrAF; and (b) provide a further specified (and standardised) vocabulary to encode these particular types of annotations. Thus, in general, these are the ones that should be used for mopho-syntactic and syntactic annotation; nevertheless, they had never been used to encode FreeLing's outputs in a standard-compliant way. Accordingly, we used ISO/MAF (2012) and ISO/SynAF (2010) to standardise the corresponding annotations in FreeLing's results.

3.1 Standardizing FreeLing's Morpho-syntactic Annotations

There are several ways to invoke the FreeLing's web service we have created, depending on the parameter values included in the call. One of them allows invoking only its POS tagging module. The native non-standardised FreeLing's POS tagging-only output obtained for the Spanish sentence 'Mi gato se llama Tiger.' ('My cat's name is Tiger.') using such a call is shown in Example 1.

As shown in this example, FreeLing's native, basic POS tagging includes the input text of the token, its associated lemma and its POS tag. They are included a file, one token per line, one item per column. The first column contains the input text of the token; the second one, its lemma; the third one, an EAGLES (1996)-conformant POS tag[7] that includes both its grammatical category and its morphosyntactic features. Additionally, some WordNet-based sense and named entity tagging can be obtained, if some suitable options are selected. When the named entity classification option is selected, the named entity tag for the token is included in the last positions of its POS tag. When the sense tagging option is selected, the identifiers of the (Euro)WordNet synsets that might be used to sense-tag the token are included in a supplementary column, separated by blanks.

```
Mi      mi      DP1CSS
gato    gato    NCMS000
se      se      P00CN000
llama   llamar  VMIP3S0
Tiger   tiger   NP00000
.       .       Fp
```

Example 1: FreeLing's native, basic POS tagging of 'Mi gato se llama Tiger.' ('My cat's name is Tiger.')

[7] For instance, the POS tag for 'Mi' ('My') is 'DP1CSS', which means that [A] it is a token whose grammatical category is determinant or pronoun ('DP'); and [B] it has the following morphosyntactic features: first person ('1'), common gender ('C'), singular number ('S') and possessive type (final 'S')).

```
grup-verb/top/s1_t13 (llama llamar VMIP3S0)s1_nt13 top_[
s1_nt13 [
    morfema-verbal/es/s1_t10 (se se P00CN000)
    sn/subj/s1_t7 (gato gato NCMS000)s1_nt7 subj_[
s1_nt7 [
        espec-ms/espec/s1_t4 (Mi mi DP1CSS)
  ]
    sn/dobj/s1_t17 (Tiger tiger NP00000)
    F-term/term/s1_t19 (. . Fp)
]
```

Example 2: FreeLing's native dependency parsing of 'Mi gato se llama Tiger.' ('My cat's name is Tiger.')

The main disadvantage of this format of annotation is that it does not state explicitly what these fields mean. Therefore, the semantics of each tag is implicit. This makes it extremely difficult (1) to make the resulting annotations be automatically interpreted and, hence, also (2) to compare them with other POS annotations and, in general, (3) to make them interoperate with other annotations (Pareja-Lora, 2012). Besides, the tokens and their POS annotations cannot be reused and/or referenced, e.g., by a syntactic annotation of the sentence, since no way to link to them is provided.

These problems were partially solved by means of the standardisation of these annotations. The standardisation of FreeLing's morphosyntactic annotations was performed according to and complying with ISO/MAF (2012) [8] . The ISO/MAF (2012)-compliant XML annotations of the Spanish sentence 'Mi gato se llama Tiger.', obtained by means of the FreeLing web service we have implemented, has been included in Example 3 (see next page).

As shown in this example, first, each `token` element is assigned a persistent identifier (PISA: its corresponding URI[9] within the file, see ISO/PISA (2011)), by means of the `@xml:id` attribute, which helps building other annotations on top of this one and linking them together (for example, the `wordForm` annotations). In this way, tokens can be referenced internally (from inside the file), locally (from inside the same [file] system) and globally (from outside the system). For instance, a token can be easily referenced locally by concatenating the identifier of the annotation file where it is included with the token identifier (see an example in Footnote 16).

Second, in order to ease the recoverability of the input text, we used the `@join` standard attribute of token elements (value: "left") to signal those cases in which no space separated two tokens (for example, within contractions).

Finally, a `wordForm` element is attached to each token[10], in order to annotate it and make the semantics of each of its tags explicit. This is achieved by means of (1) the `@lemma` attribute of wordforms and (2) a nested standard-compliant feature structure annotation (ISO/FSR (2006)) element (`fs`), which encapsulates the rest of its features[11].

3.2 Standardizing FreeLing's Syntactic Annotations

FreeLing provides both constituency-based and dependency-based syntactic annotations of its inputs. Both types of annotations have been standardized already in the web service. However, in this section we will refer mainly to its dependency-based annotations for the sake of space.

The native, non-standardised FreeLing's dependency parsing of the Spanish sentence 'Mi gato se llama Tiger.' ('My cat's name is Tiger.') is shown in Example 2 (above). As shown in this example, FreeLing's native dependency-based parser (A) uses its own parenthetical and non-semantically explicit notation to encode syntactic annotations; and (B) includes also a POS tagging of the input [12]. A graphical representation of this dependency parsing (a screenshot of FreeLing's online demo[13]) has been included for clarity in Figure 1 (after Example 3).

On the one hand, (A) clearly complicates interpreting the annotations and making them interoperate; on the other hand, regarding (B), even though having both morphosyntactic and syntactic annotations together helps making them interoperate, neither FreeLing's native dependency-based (or its constituency-based) annotations nor their nested morphosyntactic annotations can be referenced from other annotations (e.g. sense tagging).

Therefore, we decided to standardise morphosyntactic and syntactic annotations separately and interlink them together afterwards[14].

---

[8] ISO/MAF (2012) provides a general framework and a set of recommendations for the annotation of morphosyntactic units with their grammatical category and its morphosyntactic features. It provides also a recommended (not mandatory) XML serialisation for morphosyntactic annotations, which makes them be more syntactically interoperable and referenceable by other annotations.

[9]   Uniform   Resource   Identifier,   see http://www.w3.org/TR/uri-clarification/.

[10] By means of the standard `@tokens` attribute.

[11] Note that wordforms are assigned their own persistent identifier (by means of the `@xml:id` attribute) as well.

[12] Both (A) and (B) hold also for FreeLing's native constituency-based annotations.

[13] http://nlp.lsi.upc.edu/freeling/demo/demo.php.

[14] Following the best practices and recommendations discussed in Pareja-Lora (2012).

```xml
<?xml version="1.0" encoding="UTF-8"?>

<maf>
   <token xml:id="t1">Mi</token>
   <token xml:id="t2">gato</token>
   <token xml:id="t3">se</token>
   <token xml:id="t4">llama</token>
   <token xml:id="t5">Tiger</token>
   <token xml:id="t6" join="left">.</token>

   <wordForm xml:id="wordForm1" tokens="#t1" lemma="mi">
      <fs>
         <f name="pos">
            <symbol value="DP1CSS"/>
                <!--Determinante: Posesivo, Primera persona, Común, Número singular, Poseedor singular-->
         </f>
      </fs>
   </wordForm>

   <wordForm xml:id="wordForm2" tokens="#t2" lemma="gato">
      <fs>
         <f name="pos">
            <symbol value="NCMS000"/>
                <!-- Nombre: Común, Masculino, Singular-->
         </f>
      </fs>
   </wordForm>

   <wordForm xml:id="wordForm3" tokens="#t3" lemma="se">
      <fs>
         <f name="pos">
            <symbol value="P00CN000"/>
                <!--Pronombre: Común, Impersonal/Invariable -->
         </f>
      </fs>
   </wordForm>

   <wordForm xml:id="wordForm4" tokens="#t4" lemma="llamar">
      <fs>
         <f name="pos">
            <symbol value="VMIP3S0"/>
                <!--Verbo: Principal, Indicativo, Presente, Tercera persona, Singular -->
         </f>
      </fs>
   </wordForm>

   <wordForm xml:id="wordForm5" tokens="#t5" lemma="tiger">
      <fs>
         <f name="pos">
            <symbol value="NP00000"/>
                <!-- Nombre: Propio, Genero Indeterminado, Numero Indeterminado-->
         </f>
      </fs>
   </wordForm>

   <wordForm xml:id="wordForm6" tokens="#t6" lemma=".">
      <fs>
         <f name="pos">
            <symbol value="Fp"/>
                <!--Puntuación: Punto Final-->
         </f>
      </fs>
   </wordForm>

</maf>
```

Example 3: MAF-compliant annotation of 'Mi gato se llama Tiger.', obtained with the FreeLing web service

func: top
synt: grup-verb
form: llama
lemma: llamar
tag: VMIP3S0

func: es
synt: morfema-verbal
form: se
lemma: se
tag: P00CN000

func: subj
synt: sn
form: gato
lemma: gato
tag: NCMS000

func: dobj
synt: sn
form: Tiger
lemma: tiger
tag: NP00000

func: term
synt: F-term
form: .
lemma: .
tag: Fp

func: espec
synt: espec-ms
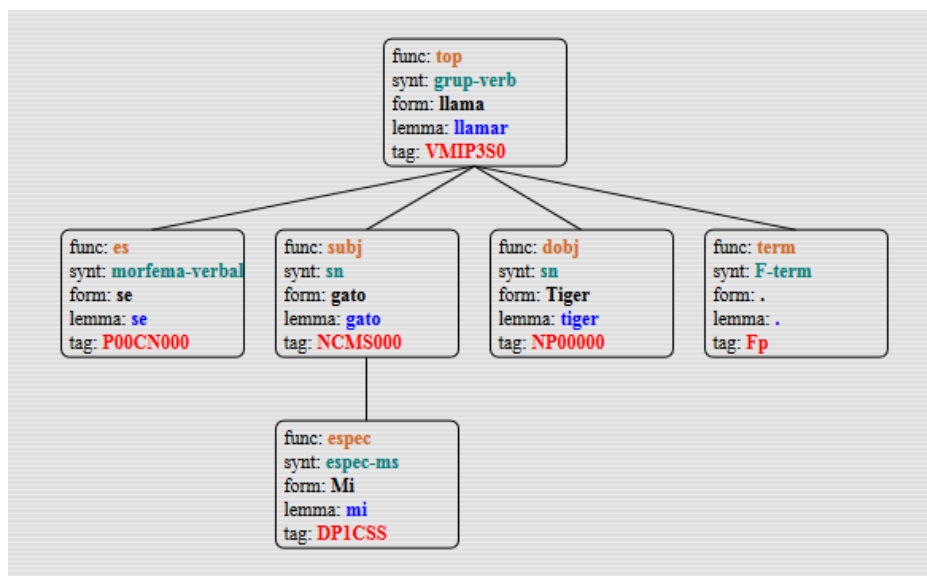form: Mi
lemma: mi
tag: DP1CSS

Figure 1: FreeLing's graphical output of the dependency parsing of 'Mi gato se llama Tiger.' ('My cat's name is Tiger.')

Hence, whereas the standardisation of FreeLing's syntactic annotations was performed according to and complying with ISO/MAF (2012), the standardisation of FreeLing's syntactic annotations was performed according to and complying with ISO/SynAF (2010), and using the XML schema included in the ISO/SynAF-<tiger2/> standard proposal, presented in Bosch et al. (2012) [15]. This twofold (and separate) standardisation also helped us test the interoperability of both ISO/MAF (2012) and ISO/SynAF (2010) compliant annotations.

The ISO/SynAF (2010)-compliant XML annotation of the Spanish sentence 'Mi gato se llama Tiger.', obtained by means of the FreeLing web service we have implemented, is shown in Example 4 (see next page). It cannot be fully described here for the sake of space; however, it is important to note that (1) the dependencies are represented by means of the <edge> elements attached to the terminal nodes (the <t> elements) and their standard @tiger2:target attribute; and (2) the terminal nodes refer to the morphosyntactic wordForm elements by means of their standard @tiger2:corresp attribute [16] being assigned the PISA of the wordforms as value.

4. Conclusions

In this paper, we have presented the transformation of the FreeLing 3.0 annotation tool into a standardised web service (FreeLing SWS). This transformation has helped us solve several interoperability limitations that FreeLing 3.0 and its morphosyntactic and syntactic annotations have (for instance, the low reusability of the tool 'as is' in NLP pipelines, and the lack of semantic explicitness and (inter-)referenceability of its annotations), as discussed in Section 3. In particular, it has helped us interlink successfully and fairly straightforwardly FreeLing's POS annotations with (i) its constituency-based syntactic annotations; and (ii) its dependency-based syntactic annotations; and make them interoperate.

However, on the one hand, unlike EAGLES (1996), ISO/MAF (2012) fails to specify in detail (i) which morphosyntactic tags are mandatory and/or recommended for each language; and (ii) the way in which a given set of morphosyntactic categories should be encoded in a particular morphosyntactic annotation scheme. Therefore, even though it helps make morphosyntactic annotations more syntactically interoperable, it does not help make them fully syntactically and/or semantically interoperable. In effect, comparing FreeLing's POS standardised annotations with other POS [standardised] annotations might require an intermediate process of tag mapping and/or translation. This is also due to the fact that annotating separately each morphosyntactic feature (e.g., grammatical category, gender, number, etc.) is not mandatory in ISO/MAF (2012). This prevents the current version of the FreeLing web service from being fully interoperable at this level. This will be tackled in a forthcoming version of FreeLing SWS.

On the other hand, even though ISO/SynAF (2010) allows interlinking constituency- and dependency-based annotations and helps making them interoperate, the automatic interlinking of these types of annotations in this standardised version of FreeLing has not been achieved so far. It will require further research, since establishing the mappings between the nodes in both annotations is not a trivial task. Therefore, this interoperability test will be tackled in a forthcoming version of the web service.

Some additional interoperability tests needed include integrating this new web service and its outputs in an actual NLP pipeline, in order to find out their real degree of interoperability. We plan, for example, to compare the efforts and resources required to integrate FreeLing SWS with those required when integrating other similar tools.

---

[15] Neither ISO/SynAF (2010) nor ISO/SynAF-<tiger2/> are introduced here for the sake of space.

[16] For example, "spanish.example.maf.xml#wordForm1" is a dereferenceable persistent identifier for the "wordForm1" ('Mi' – 'My'), of the "spanish.example.maf.xml" local file.

```xml
<?xml version="1.0" encoding="UTF-8" standalone="true"?>

<corpus    xsi:schemaLocation="http://korpling.german.hu-berlin.de/tiger2/V2.0.5/
                          http://korpling.german.hu-berlin.de/tiger2/V2.0.5/Tiger2.xsd"
           xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
           xmlns:tiger2="http://korpling.german.hu-berlin.de/tiger2/V2.0.5/"
           xmlns="http://korpling.german.hu-berlin.de/tiger2/V2.0.5/">

    <head>
        <meta>
            <name>http://quijote.fdi.ucm.es:8084/FreeLingWebService/2014-12-21.0:08:36./
                 example_tiger_dep.standoff.tiger2.xml</name>
            <author>FreeLing SWS</author>
            <date>2014/03/21</date>
            <description>Tiger2 XML syntactic dependency annotations (SynAF-2 compliant)</description>
            <format>FreeLing Tagset</format>
            <history>version:3.0</history>
        </meta>

        <annotation>
            <external
            corresp="http://quijote.fdi.ucm.es:8084/FreeLingWebService/tagsets/MyAnnotations.xml"/>
        </annotation>
    </head>

    <body>
      <s xml:id="s1">
        <graph discontinuous="false" root="s1_Root">
            <terminals>

                <t  tiger2:corresp="http://quijote.fdi.ucm.es:8084/FreeLingWebService/
                                 2014-12-21.0:08:36./spanish.example.MAF_dep.maf.xml#wordForm1"
                                 xml:id="s1_t4" synt-type="espec-ms">
                    <!-- Mi -->
                    <edge tiger2:target="s1_t7" label="espec" tiger2:type="prim.dep"/>
                </t>

                <t tiger2:corresp="http://quijote.fdi.ucm.es:8084/FreeLingWebService/
                                 2014-12-21.0:08:36./spanish.example.MAF_dep.maf.xml#wordForm2"
                                 xml:id="s1_t7" synt-type="sn">
                    <!-- gato -->
                    <edge tiger2:target="s1_t13" label="subj" tiger2:type="prim.dep"/>
                </t>

                <t tiger2:corresp="http://quijote.fdi.ucm.es:8084/FreeLingWebService/
                                 2014-12-21.0:08:36./spanish.example.MAF_dep.maf.xml#wordForm3"
                                 xml:id="s1_t10" synt-type="morfema-verbal">
                    <!-- se -->
                    <edge tiger2:target="s1_t13" label="es" tiger2:type="prim.dep"/>
                </t>

                <t tiger2:corresp="http://quijote.fdi.ucm.es:8084/FreeLingWebService/
                                 2014-12-21.0:08:36./spanish.example.MAF_dep.maf.xml#wordForm4"
                                 xml:id="s1_t13" synt-type="grup-verb">
                    <!-- llama -->
                    <edge tiger2:target="s1_Root" label="top" tiger2:type="prim.dep"/>
                </t>

                <t tiger2:corresp="http://quijote.fdi.ucm.es:8084/FreeLingWebService/
                                 2014-12-21.0:08:36./spanish.example.MAF_dep.maf.xml#wordForm5"
                                 xml:id="s1_t17" synt-type="sn">
                    <!-- Tiger -->
                    <edge tiger2:target="s1_t13" label="dobj" tiger2:type="prim.dep"/>
                </t>

                <t tiger2:corresp="http://quijote.fdi.ucm.es:8084/FreeLingWebService/
                                 2014-12-21.0:08:36./spanish.example.MAF_dep.maf.xml#wordForm6"
                                 xml:id="s1_t19" synt-type="F-term">
                    <!-- . -->
                    <edge tiger2:target="s1_t13" label="term" tiger2:type="prim.dep"/>
                </t>
            </terminals>
        </graph>
      </s>
    </body>
```

Example 4: FreeLing SWS's SynAF- and SynAF-<tiger2/>-compliant annotation of 'Mi gato se llama Tiger.'

5.   Acknowledgements

6.   References

Ballesteros-Calvo, A., Cárcamo-Escorza, G. & Duobert-Collazos, E. (2013). Recubrimiento y normalización de recursos lingüísticos: aplicación a las anotaciones morfosintácticas y sintácticas de FreeLing. B.Sc. Degree. Facultad de Informática, Universidad Complutense de Madrid. Madrid, Spain.

Bosch, S., Choi, K.S., de La Clergerie, É., Fang, A.C., Faass, G., Lee, K., Pareja-Lora, A., Romary, L., Witt, A., Zeldes, A. & Zipser, F. (2012). <tiger2/> as a Standardised Serialisation for ISO 24615 – SynAF. In Proceedings of the Eleventh International Workshop on Treebanks and Linguistic Theories (TLT11). Lisboa: Ediçôes Colibri, pp. 37 - 60.

Chinchor, N. (1997). MUC-7 Named entity Task Definition Version 3.5. [Avail. online at http://itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/ne_task.html (visited on 23/10/2013)].

EAGLES Consortium (1996). EAGLES: Recommendations for the Morphosyntactic Annotation of Corpora. EUROPEAN PROJECT DELIVERABLE: EAGLES Document EAG--TCWG—MAC/R. [Available online at http://www.ilc.cnr.it/EAGLES96/annotate/annotate.html (visited on 17/01/2012)].

Fellbaum, C. (1998, ed.). WordNet: An Electronic Lexical Database. Cambridge, MA: MIT Press.

Ide, N. & Suderman, K. (2007). GrAF: A Graph-based Format for Linguistic Annotations. In Proceedings of the Linguistic Annotation Workshop, held in conjunction with ACL'2007. Prague, Czech Republic. June, 2007, pp. 1-8.

International Organization for Standardization (2006). ISO 24610-1:2006. Language resource management -- Feature structures -- Part 1: Feature structure representation.

International Organization for Standardization (2010). ISO 24615:2010. Language resource management – Syntactic annotation framework (SynAF).

International Organization for Standardization (2011). ISO 24619:2011. Language resource management – Persistent identification and sustainable access (PISA).

International Organization for Standardization (2012). ISO 24611:2012. Language resource management – Morpho-syntactic annotation framework (MAF).

International Organization for Standardization (2012). ISO 24612:2012.Language resource management – Linguistic annotation framework (LAF).

Kashyap, V., Bussler, C. & Moran, M. (2008). The Semantic Web. Berlin: Springer.

König, E., Lezius, W. (2003). The TIGER language - A Description Language for Syntax Graphs, Formal Definition. Technical report IMS, Universität Stuttgart, Germany.

Miller, G.A. (1995). WordNet: A Lexical Database for English. In Communications of the ACM, Vol. 38, No. 11: 39-41.

Morell, C., Vivaldi, J. & Bel, N. (2012). "Iula2Standoff: a tool for creating standoff documents for the IULACT". In Calzolari, N., (et al.) (eds.) Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'12). Paris: European Language Resources Association (ELRA).

Padró, L., Stanilovsky, E. (2012). FreeLing 3.0: Towards Wider Multilinguality. Proceedings of the 8th Language Resources and Evaluation Conference (LREC 2012). ELRA. Istanbul, Turkey. May, 2012.

Pareja-Lora, A. (2012). Providing Linked Linguistic and Semantic Web Annotations: The OntoTag Hybrid Annotation Model. Saarbrücken: LAP - LAMBERT Academic Publishing, pp. 1-500.

Poch, M. & Bel, Núria. (2011). Interoperability and technology for a language resources factory. In IJCNLP 2011 Workshop on Language Resources, Technology and Services in the Sharing Paradigm. Chiang Mai, Thailand. November, 2011.

Vossen, P. (1998, ed.). EuroWordNet: a multilingual database with lexical semantic networks. Norwell, MA: Kluwer Academic Publishers.

---

[17] A **S**ocial and **O**ntology-based framework for **C**ognitively-**A**ugmented **L**anguage **L**earning in **M**obile **E**nvironments (SO-CALL-ME).