10.12871/CLICIT2014130

A Reevaluation of Dependency Structure Evaluation

Rodolfo Delmonte
Dipartimento di Studi Linguistici e Culturali Comparati
Ca' Bembo – Dorsoduro 1075
Università Ca' Foscari – 30123 VENEZIA
Email: delmont@unive.it

Abstract

English. In this paper we will develop the argument indirectly raised by the organizer of 2014 Dependency Parsing for Information Extraction task when they classify 19 relations out of 45 as those semantically relevant for the evaluation, and exclude the others which confirms our stance which considers the new paradigm of Dependency parsing evaluation favoured in comparison to the previous parsing scheme based mainly on constituent or phrase structure evaluation. We will also speak in favour of rule-based dependency parsing and against statistically based dependency parsers for reasons related to the role played by the SUBJect relation in Italian.

Italiano. *In questo lavoro svilupperemo un argomento indirettamente sollevato dagli organizzatori del task "2014 Dependency Parsing for Information Extraction", quando classificano 19 relazione come semanticamente rilevanti delle 45 presenti ed escludono le altre. Questo conferma la nostra posizione che considera il paradigma della valutazione dei parser a dipendenze favorito se confrontato con il precedente schema di parsing basato principalmente sulla valutazione della costituenza o strutture sintagmatiche. Parleremo anche a favore del parsing a dipendenze basato su regole e contro i parser a dipendenze solo statistici per ragioni relative al ruolo giocato dal ruolo di SOGGetto in italiano.*

1 Introduction

In this paper I will question the currently widely spread assumption that Dependency Structures (hence DS) are the most convenient syntactic representation, when compared to phrase or constituent structure. I will also claim that evaluation metrics applied to DS are somehow "boasting" its performance with respect to phrase structure (hence PS) representation, without a real advantage, or at least it has not yet been proven there is one. In fact, one first verification has been achieved by this year Evalita Campaign which has introduced a new way of evaluating Dependency Structures, called DS for Information Extraction - and we will comment on that below[1].

In the paper I will also argue that some features of current statistical dependency parsers speak against the use of such an approach to the parsing of languages like Italian which have a high percentage of non-canonical structures (hence NC). In particular I will focus on problems raised by the way in which SUBJect arguments are encoded. State of the art systems are using more and more dependency representations which have lately shown great resiliency, robustness, scalability and great adaptability for semantic enrichment and processing. However, by far the majority of systems available off the shelf don't support a fully semantically consistent representation and lack Empty or Null Elements (see Cai et al. 2001)[2].

O.Rambow (2010) in his opinion paper on the relations between dependency and phrase structure representation has omitted to mention the most important feature that differentiates them. PS evaluation is done on the basis of Brackets, where each bracket contains at least one HEAD, but it may contain other Heads nested inside. Of course, it may also contain a certain number of minor categories which however don't count for evaluation purposes. On

[1] As we read in the details of the call published on the Evalita website:

> "The output of participant systems will be evaluated on the basis of two scoring mechanisms focusing respectively on the parsing performance and suitability for IE… In particular, evaluation will focus on a selection of relations (19 out of a total of 45) chosen according to the following general criteria:
> a. semantic relevance of the relation (i.e. nsubj, dobj ...)
> b. exclusion of syntactic easy to identify relations (i.e. det, aux ...);
> c. exclusion of sparse and difficult to identify relations (i.e. csubj)"

[2] Additional problems are raised by the existence of Non-projective relations which amount to a consistent number of displaced constituents, both as Arguments and as Adjuncts, as discussed below.

the contrary, DS evaluation is done on the basis of head-dependent relations intervening between a pair of TOKENs. So on the one side, F-measure evaluates number of brackets which coincide with number of Heads; on the other side it evaluates number of TOKENS. Now, the difference in performance is clearly shown by percent accuracy obtained with PS evaluation which for Italian was contained in a range between 70% to 75% in Evalita 2007, and between 75% and 80% in Evalita 2009 – I don't take into account 2011 results which are referred to only one participant. DS evaluation reached peaks of 95% for UAS and in between 84% and 91% for LAS evaluation. Since data were the same for the two campaigns, one wonders what makes one representation more successful than the other.

Typically, constituent parsing is evaluated on the basis of constituents, which are made up of a head and an internal sequence of minor constituents dependent on the head. What is really important in the evaluation is the head of each constituent and the way in which PS are organized, and this corresponds to bracketing. On the contrary, DS are organized on the basis of a "word level grammar", so that each TOKEN constributes to the overall evaluation, including punctuation (not always). Since minor categories are by far the great majority of the tokens making up a sentence – in Western languages, but no so in Chinese, for instance (see Yang & Xue, 2010)– the evaluation is basically made on the ability of the parser to connect minor categories to their heads.

What speaks in favour of adopting DS is the clear advantage gained in the much richer number of labeled relations which intervene at word level, when compared to the number of constituent labels used to annotate PS relations[3]. It is worth while noting that DS is not only a much richer representation than PS, but it encompasses different levels of linguistic knowledge. For instance, punctuation may be used to indicate appositions, parentheticals, coordinated sets, elliptical material, subdivision of complex sentences into main and subordinate clause. The same applies to discourse markers which may be the ROOT of a sentence. These have all to be taken into account when computing DS but not with PS parsing.

2 Hypotheses about Dependency Evaluation Success Story

In every Western language, the number of SEMANTIC heads – Nouns, Verbs, Adjectives and Adverbs – is very low when compared to the number of tokens. Rank lists for Italian and English in their upper part are cluttered with articles, prepositions, conjunctions, quantifiers and other determiners. Semantically relevant words only come below a certain frequency threshold. To ascertain these proportions, we decided to look into the dependency treebank made freely available for the current Evalita campaign: here below is a statistics of heads which play the role of Arguments and then those that play both the role of Argument and that of Adjuncts. Percentages are obtained by dividing each relation total occurrency with the total number of tokens, which is 158485.

As can be easily noticed, Core Arguments only make 10% of all tokens and even in a 90% accuracy test result all of them might be wrong. Notice that NSUBJs include 1049 NCUSBJPASS.

	Occur.	**Percent**
nsubj	7549	4.5518%
dobj	5519	3.3278%
iobj	852	0.5137%
xcomp	1036	0.6242%
acomp	1020	0.6150%
TotCore	***15976***	***10.008%***
pobj	23313	14.058%
TotalC+P	***39289***	***24.79%***
csubj	187	
vmod	7920	
rcmod	1945	
TotalAdj	***10052***	***6.343%***
ROOT	7399	
TotalC+P+A+R	***56740***	***35.801%***

Table 1. Grammatical Relations in SIDT

POBJ include both Oblique arguments – a small part - and circumstantial Adjuncts – the great majority. We know for sure that Oblique arguments usually occur with intransitive verbs, which are a small percentage of all verbs. They may also occur as arguments of Ditransitive verbs, but also these are a small percentage of Italian verbs. So we may well say that core Argument grammatical relations only cover some

[3] In particular, then, there is at least one relation lacking in PS representation and is coordination, which on contrary is always represented in DS. As for grammatical relations like SUBJect and OBJect, they are usually not available in PS but they actually appear in PennTreebank II and so they can be learned.

12% of all tokens. Considering that the 64% of all tokens have minor or secondary dependency relations – those based on minor categories or punctuation -, we come up with the conclusion that the remaining 27% needed to cover the best accuracy result obtained so far (91%) is scattered amongst Arguments and Adjuncts. But Arguments and Adjuncts head-dependent relations only constitute 36% of all dependency relations and 27% will only cover 75% of them, no more. So eventually, an evaluation based on semantically relevant heads of Arguments and Adjuncts will achieve worse results than one based on phrase structures.

Now consider ROOT heads which include also root heads of fragments, typically nominal heads. The total number of Inflected verbs in the treebank amounts to 10800 heads. This means that the percentage of null subject elements is 30.102% of all inflected clauses - we subtract expressed subjects from total inflected verbs 10800-7549 = 3251. This 30% of missing SUBJect arguments deteriorates any evaluation. Then we need to consider that there will be another 30% of subjects which are difficult to get because they are placed in noncanonical position – this is derived from a statistics based on VIT(see Delmonte et al. 2007)[4]. It is a fact that in this way, the semantics of the representation used and produced at runtime becomes inconsistent and will reduce dramatically its usefulness in real life applications like Information Extraction, Q/A and other semantically driven fields by hampering the mapping of a complete logical form. Statistical models for DS only encode lexically expressed subjects, null elements being strictly forbidden.

Coming now to general results of the Relations Task in Evalita - specific results are not yet available -, we see that Precision best percentage almost reaches 82%, while the worst is around 78%. However seen that Recall is in the range of 90-85% the accuracy would average 80%. F1 is subsequently contained in the range 86-83%. Data are then equal to if not worse than those of PS evaluation. Even though we don't have available a detailed distribution of the data in the different categories, we may definitely say that they confirm our stance. Thus we expect minor categories like DET to be correct at 98%; not so for those relevant relations corresponding to semantic heads.

2.1 Problematic issues for statistical parsers of Italian

There are two types of Dependency parsers: rule-based symbolic parsers which can also make use of statistics at some step of computation; and statistically only parsers which make use of a classifier and a model to decide how to process the input word (see Delmonte, 1999; 2000; 2002; 2005). The second one could also be – as is the case of Stanford parser (see De Marneff et al., 2011) - a phrase structure probabilistic parser with a mapping or conversion step of syntactic constituents into DS.

Statistical dependency parsers are trained on annotated treebank data and make predictions on the basis of the model. They tap their knowledge from a training corpus which leads to the creation of a model using a classifier. The fundamental idea is the ability of the parser to use the model in a predictive way in order to generalize the encoded information to new and unseen linguistic material[5]. Even if it is obvious that a statistical model can represent linguistic knowledge at any depth and level of representation by increasing the number of features, this is not always convenient both on grounds of efficiency and overall performance. However, linguistic knowledge is split into two main components: the grammar and the lexicon. It is reasonable to assume that learning can only be achieved for the grammatical component and only for regular linguistic phenomena. The other important component, the lexicon, is on the contrary not predictable by definition. Lexical knowledge is idiosyncratic and unpredictable: for instance,

[4] We verified the proportion of null subject in VIT is even higher. We derived data of non-canonical structures in Italian from the treebank called VIT - Venice Italian Treebank (see Delmonte et al., 2007; Delmonte 2009) and compared them to PennTreebank data. In particular, VIT has 36.46% of NC structures vs. 13.01% for PT. As to lexically unexpressed or null subject, VIT has 51.31% vs 0.27% for PT. NC structures are measured against the number of total utterances that is 10,200 for VIT, and 55,600 for PT. On the contrary, Null Subjects are counted against simple sentences, that is 19,099 for VIT and 93,532 for PT. As for Subjects, there were 6230 canonical - i.e. strictly in preverbal position with no linguistic material intervening between Subject and inflected Verb - lexically expressed SUBJects out of the total 10,100 lexically expressed SUBJects. This means that non-canonical subjects constitutes 1/3 of all expressed SUBJects.

[5] In fact, when used in a different domain, the same parser is usually susceptible to serious performance degradation which can get as high as 14% (see Lease & Charniak, 2004). This problem has been partially solved by introducing several parser adaptation techniques to new domains.

knowing that certain verbs belong to the class of atmospheric or impersonal verbs and are associated to special constructions simply requires knowing which they are. Grammatical knowledge is on the contrary predictable being associated either to grammatical or functional words - which are very frequent, or to the presence of specific morphemes[6].

Given the great variety of possible structures in Italian sentences, it is quite reasonable to assume that they may suffer from problems related to the SUBJect relation. Parsers of Italian are in general unable to detect duplicate subjects, and can erroneously licence a proposition or clause without a subject even if one is available but not in canonical position. Since Italian has null subjects, this may happen quite frequently. Just for the matter of documenting the phenomenon, we will show one such example below. The example is useful for two reasons: it shows two different approaches to parsing (one without and one with null elements); secondly it helps documenting the phenomenon.

We take Null subject in Italian to be a feature that speaks in favour of rule-based parsers. Rule-based parser have more resiliency and don't need any training. They can base their knowledge on the lexicon where selectional preferences are encoded, and can produce empty categories. We will use one such parser as example, and we are here referring to TULE TUT parser documented in Lombardo, Lesmo's (1998) paper. In order to show how this may affect the output representation, we report in the appendix one sentence parsed by TALN/DeSR parser (see Attardi, 2006; Attardi et al. 2009), available as webservice at http://tanl.di.unipi.it/ it/. This parser is regarded one of the best statistical dependency parsers of Italian, achieving best results in Evalita campaigns. The output is reported in Appendix 1.

(1) E dovranno riportare per ogni unità urbana anche i dati di superficie espressi in metri quadri in conformità alle istruzioni che saranno fornite in seguito, poiché questo sarà in futuro il parametro in base al quale sarà decretato l'esborso del contribuente al posto dei <vani utili> che andranno in soffitta. [7]

I used half of a very long sentence taken from an Italian administrative bill expressed in a style which is considered "burocratic" style. In DeSR output I marked with double stars all cases of wrong argument selection, and by indenting all cases of relative clauses which have no indication of argument nor grammatical relation – COMP is a generic label and should have been substituted by NSUBJ, DOBJ or POBJ according to the grammatical relation held by the relative pronoun. Wrong argument selection in one case of double subject assignment, as well as two cases of no subject assignment. Errors may be partly due to wrong tagging disambiguation.

The same sentence has been parsed by TULETUT parser which is able to process the two relative structures with almost no error. One of the reasons for this difference, maybe because it uses subcategorization information. In addition, TULETUT parser also correctly produces empty subject categories and traces for long distance dependencies. However, also this representation has one error, and it is the missing link between the relative pronoun and its governing verb: as it is usually the case, the relative pronoun is linked to the verb of the internal clause and the verb is linked to antecedent. This does not happen with relative prepositional object IN_BASE_AL QUALE.

3 Conclusion

In this paper I tried to highlight critical issues on the current way of evaluating DS which indirectly "boasts" the performance of the parsers when compared to phrase structure evaluation. I assume this is due to the inherent shortcoming of DS evaluation not considering semantically relevant grammatical relations as being more important than minor categories. Statistical dependency parsers may have more problems in encoding features of Italian Subject because of its multiple free representations. For this reasons, I argued in favour of rule-based dependency parsers and I presented in particular, one example from TULETUT, a deep parser of Italian.

[6] In fact, for some linguistic theories - the constructional theory of grammar - also syntactic constructions are part of lexical knowledge (see Goldberg, 2006).

[7] This is a literal translation: "And they should include for each urban unit also surface data expressed in square meters in compliance with the instructions that will be made available later, because this will be in future the parameter on the basis of which will be decided the payment by the tax-payers in place of <useful rooms> which will be abandoned."

References

Attardi, G. 2006. Experiments with a Multilanguage Non-Projective Dependency Parser. Proc. of the Tenth Conference on Natural Language Learning, New York, (NY), 2006.

Attardi, G., F. Dell'Orletta, M. Simi, J. Turian. 2009. Accurate Dependency Parsing with a Stacked Multilayer Perceptron. Proc. of Workshop Evalita.

Cai, Shu, David Chiang, Yoav Goldberg, 2011. Language-Independent Parsing with Empty Elements, in Proceedings of the 49th Annual Meeting of the ACL, 212–216.

De Marneffe, M.C., B. MacCartney, C. D. Manning, Generating typed dependency parses from phrase structure parses, Proceedings of LREC, 2006/5, 449-454.

Delmonte R. 1999. From Shallow Parsing to Functional Structure, in Atti del Workshop AI*IA - "Elaborazione del Linguaggio e Riconoscimento del Parlato", IRST Trento, pp.8-19.

Delmonte R. 2000. Parsing Preferences and Linguistic Strategies, in LDV-Forum - Zeitschrift fuer Computerlinguistik und Sprachtechnologie - "Communicating Agents", Band 17, 1,2, pp. 56-73.

Delmonte R. 2002. GETARUN PARSER - A parser equipped with Quantifier Raising and Anaphoric Binding based on LFG, Proc. LFG2002 Conference, Athens, pp.130-153, at http://cslipublications.stanford.edu/hand/miscpubs online.html.

Delmonte R., 2005, Deep & Shallow Linguistically Based Parsing, in A.M.Di Sciullo(ed), UG and External Systems, John Benjamins, Amsterdam/Philadelphia, pp.335-374.

Delmonte, R., A. Bristot, S. Tonelli 2007. VIT - Venice Italian Treebank: Syntactic and Quantitative Features, in K. De Smedt, Jan Hajic, Sandra Kübler(Eds.), Proc. Sixth International Workshop on TLT, Nealt Proc. Series Vol.1, 43-54.

Delmonte R., 2009. Treebanking in VIT: from Phrase Structure to Dependency Representation, in Sergei Nirenburg (ed.), Language Engineering for Lesser-Studied Languages, IOS Press, Amsterdam, The Netherlands, 51-80.

Goldberg, A. E. 2006. Constructions at work: The nature of generalization in language. Oxford: Oxford University Press.

Lease M. & Eugene Charniak, 2005. Parsing biomedical literature. in Proceedings IJCNLP ’05, Second International Joint Conference, Jeju Island, Korea. pp 58-69.

Lombardo, V. & L. Lesmo, 1998. Formal aspects and parsing issues of dependency theory. In Coling-ACL ’98, 787-793, Montreal, Canada.

Rambow, Owen (2010). The simple truth about dependency and phrase structure representations: an opinion piece. In Proceedings of NAACL-HLT, 1-4.

Yang, Yaqin and Nianwen Xue. 2010. Chasing the ghost: recovering empty categories in the Chinese Treebank.In Proc. COLING.

APPENDIX 1. TALN/DeSR parser output of sentence (1)

ID	FORM	LEMMA	CPOS	POS	MORPH	HEAD	DEP	PH	PD
1	e	e	C	CC	_	0	ROOT	_	_
2	dovranno	dovere	V	VM	num=p|per=3|mod=i|ten=f	3	aux	_	_
3	riportare	riportare	V	V	mod=f	1	conj	_	_
4	per	per	E	E	_	3	comp	_	_
5	ogni	ogni	D	DI	num=s|gen=n	6	mod	_	_
6	unità	unità	S	S	num=n|gen=f	4	prep	_	_
7	urbana	urbano	A	A	num=s|gen=f	6	mod	_	_
8	anche	anche	B	B	_	10	mod	_	_
9	i	il	R	RD	num=p|gen=m	10	det	_	_
10	dati	dato	S	S	num=p|gen=m	3	obj	_	_
11	di	di	E	E	_	10	comp	_	_
12	superficie	superficie	S	S	num=s|gen=f	11	prep	_	_
13	espressi	esprimere	V	V	num=p|mod=p|gen=m	10	mod	_	_
14	in	in	E	E	_	13	comp	_	_
15	metri	metro	S	S	num=p|gen=m	14	prep	_	_
**16	quadri	quadro	S	S	num=p|gen=m	13	subj_pass	_	_
17	in	in	E	E	_	13	comp	_	_
18	conformità	conformità	S	S	num=n|gen=f	17	prep	_	_
19	alle	al	E	EA	num=p|gen=f	18	comp	_	_
	20 istruzioni	istruzione	S	S	num=p|gen=f	19	prep	_	_

<table>
<tr><td>**21 che</td><td></td><td>che</td><td>P</td><td>PR</td><td>num=n|gen=n</td><td>23</td><td>comp</td><td>_</td><td>_</td></tr>
<tr><td></td><td>22 saranno</td><td>essere</td><td>V</td><td>VA</td><td>num=p|per=3|mod=i|ten=f</td><td>23</td><td>aux</td><td>_</td><td>_</td></tr>
<tr><td></td><td>23 fornite</td><td>fornire</td><td>V</td><td>V</td><td>num=p|mod=p|gen=f</td><td>20</td><td>relcl</td><td>_</td><td>_</td></tr>
<tr><td>24</td><td>in</td><td>in</td><td>E</td><td>E</td><td>_</td><td>23</td><td>comp</td><td>_</td><td>_</td></tr>
<tr><td>25</td><td>seguito</td><td>seguito</td><td>S</td><td>S</td><td>num=s|gen=m</td><td>24</td><td>prep</td><td>_</td><td>_</td></tr>
<tr><td>26</td><td>,</td><td>,</td><td>F</td><td>FF</td><td>_</td><td>3</td><td>punc</td><td>_</td><td>_</td></tr>
<tr><td>27</td><td>poiché</td><td>poiché</td><td>C</td><td>CS</td><td>_</td><td>3</td><td>mod</td><td>_</td><td>_</td></tr>
<tr><td>**28</td><td>questo</td><td>questo</td><td>P</td><td>PD</td><td>num=s|gen=m</td><td>29</td><td>subj</td><td>_</td><td>_</td></tr>
<tr><td>29</td><td>sarà</td><td>essere</td><td>V</td><td>V</td><td>num=s|per=3|mod=i|ten=f</td><td>27</td><td>sub</td><td>_</td><td>_</td></tr>
<tr><td>30</td><td>in</td><td>in</td><td>E</td><td>E</td><td>_</td><td>29</td><td>pred</td><td>_</td><td>_</td></tr>
<tr><td>31</td><td>futuro</td><td>futuro</td><td>S</td><td>S</td><td>num=s|gen=m</td><td>30</td><td>prep</td><td>_</td><td>_</td></tr>
<tr><td>32</td><td>il</td><td>il</td><td>R</td><td>RD</td><td>num=s|gen=m</td><td>33</td><td>det</td><td>_</td><td>_</td></tr>
<tr><td>**33</td><td>parametro</td><td>parametro</td><td>S</td><td>S</td><td>num=s|gen=m</td><td>39</td><td>subj</td><td>_</td><td>_</td></tr>
<tr><td></td><td>34 in</td><td>in</td><td>E</td><td>E</td><td>_</td><td>33</td><td>comp</td><td>_</td><td>_</td></tr>
<tr><td></td><td>35 base</td><td>base</td><td>S</td><td>S</td><td>num=s|gen=f</td><td>34</td><td>concat</td><td>_</td><td>_</td></tr>
<tr><td></td><td>36 al</td><td>al</td><td>E</td><td>EA</td><td>num=s|gen=m</td><td>35</td><td>concat</td><td>_</td><td>_</td></tr>
<tr><td>**37 quale</td><td></td><td>quale</td><td>P</td><td>PR</td><td>_</td><td>36</td><td>prep</td><td>_</td><td>_</td></tr>
<tr><td></td><td>38 sarà</td><td>essere</td><td>V</td><td>V</td><td>num=s|per=3|mod=i|ten=f</td><td>39</td><td>aux</td><td>_</td><td>_</td></tr>
<tr><td></td><td>39 decretato</td><td>decretare</td><td>V</td><td>V</td><td>num=s|mod=p|gen=m</td><td>29</td><td>conj</td><td>_</td><td>_</td></tr>
<tr><td>40</td><td>l'</td><td>il</td><td>R</td><td>RD</td><td>num=s|gen=n</td><td>41</td><td>det</td><td>_</td><td>_</td></tr>
<tr><td>**41</td><td>esborso</td><td>esborso</td><td>S</td><td>S</td><td>num=s|gen=m</td><td>39</td><td>subj_pass</td><td>_</td><td>_</td></tr>
<tr><td>42</td><td>del</td><td>di</td><td>E</td><td>EA</td><td>num=s|gen=m</td><td>41</td><td>comp</td><td>_</td><td>_</td></tr>
<tr><td>43</td><td>contribuente</td><td>contribuente</td><td>S</td><td>S</td><td>num=s|gen=n</td><td>42</td><td>prep</td><td>_</td><td>_</td></tr>
<tr><td>44</td><td>al</td><td>al</td><td>E</td><td>EA</td><td>num=s|gen=m</td><td>41</td><td>comp</td><td>_</td><td>_</td></tr>
<tr><td>45</td><td>posto</td><td>posto</td><td>S</td><td>S</td><td>num=s|gen=m</td><td>44</td><td>prep</td><td>_</td><td>_</td></tr>
<tr><td>46</td><td>dei</td><td>di</td><td>E</td><td>EA</td><td>num=p|gen=m</td><td>45</td><td>comp</td><td>_</td><td>_</td></tr>
<tr><td>47</td><td><</td><td><</td><td>N</td><td>N</td><td>_</td><td>48</td><td>mod</td><td>_</td><td>_</td></tr>
<tr><td></td><td>48 vani</td><td>vano</td><td>S</td><td>S</td><td>num=p|gen=m</td><td>46</td><td>prep</td><td>_</td><td>_</td></tr>
<tr><td></td><td>49 utili</td><td>utile</td><td>A</td><td>A</td><td>num=p|gen=n</td><td>50</td><td>mod</td><td>_</td><td>_</td></tr>
<tr><td></td><td>50 ></td><td>></td><td>S</td><td>S</td><td>num=p|gen=m</td><td>48</td><td>mod</td><td>_</td><td>_</td></tr>
<tr><td>**51 che</td><td></td><td>che</td><td>P</td><td>PR</td><td>num=n|gen=n</td><td>52</td><td>comp</td><td>_</td><td>_</td></tr>
<tr><td></td><td>52 andranno</td><td>andare</td><td>V</td><td>V</td><td>num=p|per=3|mod=i|ten=f</td><td>48</td><td>relcl</td><td>_</td><td>_</td></tr>
<tr><td>53</td><td>in</td><td>in</td><td>E</td><td>E</td><td>_</td><td>52</td><td>comp</td><td>_</td><td>_</td></tr>
<tr><td>54</td><td>soffitta</td><td>soffitta</td><td>S</td><td>S</td><td>num=s|gen=f</td><td>53</td><td>prep</td><td>_</td><td>_</td></tr>
<tr><td>55</td><td>.</td><td>.</td><td>F</td><td>FS</td><td>_</td><td>1</td><td>punc</td><td>_</td><td>_</td></tr>
</table>

APPENDIX 2. TULETUT parser output of sentence (1)

1 E (E CONJ COORD COORD) [0;TOP-CONJ]
2 dovranno (DOVERE VERB MOD IND FUT INTRANS 3 PL) [1;COORD2ND]
2.10 t [] (GENERIC-T PRON PERS ALLVAL ALLVAL ALLVAL) [2;VERB-SUBJ]
3 riportare (RIPORTARE VERB MAIN INFINITE PRES TRANS) [2;VERB+MODAL-INDCOMPL]
3.10 t [2.10f] (GENERIC-T PRON PERS ALLVAL ALLVAL ALLVAL) [3;VERB-SUBJ]
4 per (PER PREP MONO) [3;RMOD]
5 ogni (OGNI ADJ INDEF ALLVAL SING) [4;PREP-ARG]
6 unità (UNITà▼ NOUN COMMON F ALLVAL) [5;DET+QUANTIF-ARG]
7 urbana (URBANO ADJ QUALIF F SING) [6;ADJC+QUALIF-RMOD]
8 anche (ANCHE ADV CONCESS) [9;ADVB+CONCESS-RMOD]
9 i (IL ART DEF M PL) [3;VERB-OBJ]
10 dati (DATO NOUN COMMON M PL) [9;DET+DEF-ARG]
11 di (DI PREP MONO) [10;PREP-RMOD]
12 superficie (SUPERFICIE NOUN COMMON F SING) [11;PREP-ARG]
13 espressi (ESPRIMERE VERB MAIN PARTICIPLE PAST TRANS PL M) [10;VERB-RMOD+RELCL+REDUC]
13.10 t [] (GENERIC-T PRON PERS ALLVAL ALLVAL ALLVAL) [13;VERB-OBJ/VERB-SUBJ]
13.11 t [] (GENERIC-T PRON PERS ALLVAL ALLVAL ALLVAL) [13;VERB-SUBJ/VERB-INDCOMPL-AGENT]
14 in (IN PREP MONO) [13;RMOD]
15 metri (METRO NOUN COMMON M PL) [14;PREP-ARG]
16 quadri (QUADRO ADJ QUALIF M PL) [15;ADJC+QUALIF-RMOD]
17 in (IN PREP MONO) [13;RMOD]
18 conformità (CONFORMITà NOUN COMMON M SING) [17;PREP-ARG]
19 alle (A PREP MONO) [13;RMOD]
19.1 alle (IL ART DEF F PL) [19;PREP-ARG]
20 istruzioni (ISTRUZIONE NOUN COMMON F PL) [19.1;DET+DEF-ARG]
21 che (CHE PRON RELAT ALLVAL ALLVAL LSUBJ+LOBJ) [23;VERB-OBJ/VERB-SUBJ]
22 saranno (ESSERE VERB AUX IND FUT INTRANS 3 PL) [23;AUX]
23 fornite (FORNIRE VERB MAIN PARTICIPLE PAST TRANS PL F) [20;VERB-RMOD+RELCL]
23.10 t [] (GENERIC-T PRON PERS ALLVAL ALLVAL ALLVAL) [23;VERB-SUBJ/VERB-INDCOMPL-AGENT]
23.11 t [] (GENERIC-T PRON PERS ALLVAL ALLVAL ALLVAL) [23;VERB-INDOBJ]
24 in (IN PREP MONO) [23;RMOD]
25 seguito (SEGUITO NOUN COMMON M SING) [24;PREP-ARG]
26 , (#\, PUNCT) [23;SEPARATOR]
27 poiché (POICHé‰ CONJ SUBORD CAUS) [23;RMOD]

28 questo (QUESTO PRON DEMONS M SING LSUBJ+LOBJ+OBL) [29;VERB-SUBJ]
29 sarà (ESSERE VERB MAIN IND FUT INTRANS 3 SING) [27;CONJ-ARG]
30 in (IN PREP MONO) [29;RMOD]
31 futuro (FUTURO NOUN COMMON M SING) [30;PREP-ARG]
32 il (IL ART DEF M SING) [29;VERB-PREDCOMPL+SUBJ]
33 parametro (PARAMETRO NOUN COMMON M SING) [32;DET+DEF-ARG]
34 in (IN_BASE_A PREP POLI LOCUTION) [29;RMOD]
35 base (IN_BASE_A PREP POLI LOCUTION) [34;CONTIN+LOCUT]
36 al (IN_BASE_A PREP POLI LOCUTION) [35;CONTIN+LOCUT]
36.1 al (IL ART DEF M SING) [34;PREP-ARG]
** 37 quale (QUALE PRON RELAT ALLVAL SING 3 LSUBJ+LOBJ+OBL) [36.1;DET+DEF-ARG]
38 sarà (ESSERE VERB AUX IND FUT INTRANS 3 SING) [39;AUX]
39 decretato (DECRETARE VERB MAIN PARTICIPLE PAST TRANS SING M) [33;VERB-RMOD+RELCL]
39.10 t [] (GENERIC-T PRON PERS ALLVAL ALLVAL ALLVAL) [39;VERB-SUBJ/VERB-INDCOMPL-AGENT]
40 l' (IL ART DEF M SING) [39;VERB-OBJ/VERB-SUBJ]
41 esborso (ESBORSO NOUN COMMON M SING) [40;DET+DEF-ARG]
42 del (DI PREP MONO) [41;PREP-RMOD]
42.1 del (IL ART DEF M SING) [42;PREP-ARG]
43 contribuente (CONTRIBUENTE NOUN COMMON ALLVAL SING) [42.1;DET+DEF-ARG]
44 al (AL_POSTO_DI PREP POLI LOCUTION) [39;RMOD]
45 posto (AL_POSTO_DI PREP POLI LOCUTION) [44;CONTIN+LOCUT]
46 dei (AL_POSTO_DI PREP POLI LOCUTION) [45;CONTIN+LOCUT]
46.1 dei (IL ART DEF M PL) [44;PREP-ARG]
47 < (#\< PUNCT) [46.1;SEPARATOR]
48 vani (VANO NOUN COMMON M PL) [46.1;DET+DEF-ARG]
49 utili (UTILE ADJ QUALIF ALLVAL PL) [48;ADJC+QUALIF-RMOD]
50 > (#\> PUNCT) [48;SEPARATOR]
51 che (CHE PRON RELAT ALLVAL ALLVAL LSUBJ+LOBJ) [52;VERB-SUBJ]
52 andranno (ANDARE VERB MAIN IND FUT INTRANS 3 PL) [48;VERB-RMOD+RELCL]
53 in (IN PREP MONO) [52;VERB-INDCOMPL-LOC+TO]
54 soffitta (SOFFITTA NOUN COMMON F SING) [53;PREP-ARG]
55 . (#\. PUNCT) [1;END]