CCG Parsing and Multiword Expressions

Miryam de Lhoneux
Crowdynews, The University of Edinburgh
miryam@crowdynews.com
COST Action IC1207 PARSEME Meeting, Malta 2015, WG2 and WG4

1 Introduction

This poster presents work carried out for my MSc dissertation (de Lhoneux, 2014) at the University of Edinburgh under the supervision of Mark Steedman and Omri Abend.

The goal of our work was to find out whether or not information about Multiword Expressions (MWEs) can improve statistical parsing with Combinatory Categorial Grammar (CCG). This is interesting because information about MWEs has recently been shown to be useful for syntactic parsing and because syntactic parsing is central to NLP. Since MWEs forced a non-modular view of grammar in Linguistics theory: Syntax and the Lexicon are not separate modules in the Grammar, as argued for example in Construction Grammar (Hoffmann and Trousdale, 2013) and since CCG adopts a grammar architecture in which syntactic information is partially encoded in the Lexicon, it seemed also particularly interesting to work on MWEs within the CCG framework.

2 Background studies

Nivre and Nilsson (2004) manually created two versions of a Treebank, one in which MWE units are joined to form a token (commonly called the 'words-with-spaces' approach) and one in which they are separate. They tested whether this 'perfect MWE recognition' could help parsing accuracy. Korkontzelos and Manandhar (2010) automatically created two versions of an unannotated corpus based on a list of MWEs randomly selected. They observed a gain in parsing accuracy when the *test* data contained MWEs joined as one token. Both studies limited the types of MWEs dealt with. Two questions remained unanswered: whether or not parsing can benefit from MWE information obtained by automatic MWE recognition and whether or not the representation of MWEs as one unit in a parsing model can improve the parsing model when used with other MWE types.

3 Methodology

For clarity, an unchanged version of the model or data is called A and when a change has been made, it is called B. By training effect is meant that the parser has learnt something useful. By parsing effect is meant that collapsed test data have helped the parser in its decisions.

We worked with CCGbank with the traditional sections 01-22 for training, 00 for development and 23 for testing. The jMWE library (Finlayson and Kulkarni, 2011) was used to detect MWEs in a sentence. Different parameters of the library were used for the different experiments. MWEs were collapsed to one unit in Trees when their units did not cross constituent boundaries, as shown in Figure 1 which is transformed as in Figure 2.

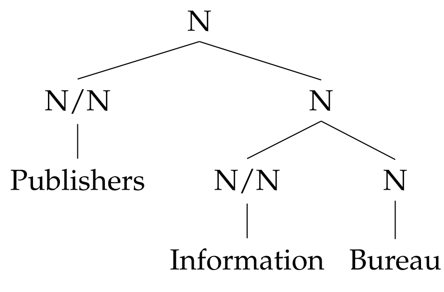

Figure 1: Original tree

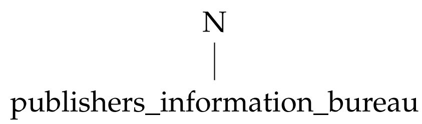

Figure 2: Collapsed tree

Dependency edges between the units of MWEs are discarded as shown in Figure 3 which is collapsed as in Figure 4.

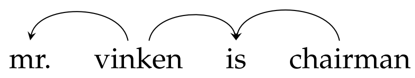

Figure 3: Original dependency graph

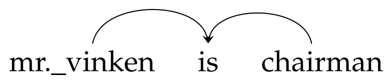

Figure 4: Collapsed dependency graph

Evaluation We computed Precision (P), recall (R) and F1 (F_1) of unlabelled dependencies of the output data against the gold standard.

To answer the question of whether or not there is a training effect, we compared the results of outA with outB (with *training* and *test* data changed) on goldB. To answer the question of wheter or not there is a parsing effect, we compared the results of outA with outB (with *test* data changed) on goldB. To find out whether or not the recognition method influence the results, we experimented and compared the results of different outB (which we decollapsed) on goldA.

4 Results

Due to a shortcoming in the methodology (the algorithm we used to modify MWEs in the Treebank is only capable of dealing with MWEs that do not cross constituent boundaries), adjustments had to be made to obtain a fair comparison of the models. This section gives result which are representative of the trends we observed.

ModelB (with both *training* and *test* data changed) was found to perform slightly but significantly (p>.05) better than ModelA on goldB which confirmed the existence of training effects, as shown in Table 1.

Data collapsed	P	R	F_1
test	**84.53**	84.76	84.64
training and test	84.48	**85.28**	**84.88**

Table 1: Training effect

ModelB (with only *test* data changed)was found to perform slightly but significantly (p>.05) better than ModelA on GoldB as shown in Table 2[1].

Data collapsed	P	R	F_1
test (before parsing)	**79.83**	79.54	**79.69**
test (after parsing)	79.38	**79.60**	79.49

Table 2: Parsing effect

Differences in results obtained with different recognizers are also small but significant (Table 3).

Data collapsed	decollapsed	MWE types handled	F_1
None	None		85.15
training and test	out	all	85.02
training and test	out	Proper Nouns	**85.28**
training and test	out	Length 2	85.07
training and test	out	Stop words	85.19

Table 3: MWE recognition experimentation

The results in Table 1 and 2 were cross-validated on goldA: by decollapsing the output of modelB, we obtained an improved performance on goldA. This improvement was however not significant.

5 Conclusion

The main contributions of our work are improvements on CCG parsing with automatic MWE recognition with significant results despite limited settings. We have developed techniques for distinguishing training from parsing effects and provided empirical support that there is both training and parsing effects. We have observed differences in results when using different recognizers

6 Future research

For future research, we suggest extending the collapsing algorithm to the non-sibling case, testing more MWE recognition methods with more data, conducting error analysis such as is done in Korkontzelos and Manandhar (2010).

Acknowledgements

Thanks are due to Mark Steedman and Omri Abend first for supervising this work but also for commenting the draft of the poster.

[1] The differences between Table 1 and Table 2 are a result of the methodology adjustments we had to make and are not meaningful.

References

de Lhoneux, M. (2014). CCG Parsing and Multiword Expressions. Master's thesis, The University of Edinburgh.

Finlayson, M. A. and Kulkarni, N. (2011). Detecting Multi-word Expressions Improves Word Sense Disambiguation. In *Proceedings of the Workshop on Multiword Expressions: From Parsing and Generation to the Real World*, MWE '11, pages 20–24, Stroudsburg, PA, USA. Association for Computational Linguistics.

Hoffmann, T. and Trousdale, G. (2013). The Handbook of Construction Grammar. In Hoffmann, T. and Trousdale, G., editors, *The Oxford Handbook of Construction Grammar*, Oxford Handbooks in Linguistics, chapter Construction Grammar: Introduction, pages 1–14. Oxford University Press.

Korkontzelos, I. and Manandhar, S. (2010). Can recognising multiword expressions improve shallow parsing? In *HLT-NAACL*, pages 636–644.

Nivre, J. and Nilsson, J. (2004). Multiword units in syntactic parsing. In *Workshop on Methodologies and Evaluation of Multiword Units in Real-World Applications*, pages 39–46.