	#gold	Stanford	DP-TSG	Berkeley
MWET	3	0.0	0.0	0.0
MWV	26	64.0	57.7	50.7
MWA	8	26.1	32.2	29.8
MWN	456	64.1	67.6	67.1
MWD	15	70.3	65.5	70.1
MWPRO	17	73.7	78.0	76.2
MWADV	220	74.6	72.7	70.4
MWP	162	81.3	80.5	77.7
MWC	47	83.5	83.5	80.8
	954	70.1	**71.1**	69.6

Table 7: MWE identification per category and overall results (test set, sentences ≤ 40 words). MWI and MWCL do not occur in the test set.

Model	F1
`mwetoolkit` All	15.4
PA-PCFG	32.6
`mwetoolkit` Filter	34.7
PA-PCFG+Features	63.1
DP-TSG	**71.1**

Table 8: MWE identification F1 of the best parsing model vs. the `mwetoolkit` baseline (test set, sentences ≤ 40 words). PA-PCFG+Features includes the grammar features in Table 4, which is the CFG from which the TSG is extracted. For `mwetoolkit`, *All* indicates the inclusion of all n-grams in the training corpus. *Filter* indicates prefiltering of the training corpus by removing rare n-grams (see §A.2 for details).

`mwetoolkit` and the CFG from the which the TSG is extracted. The TSG-based parsing model outperforms `mwetoolkit` by 36.4% F1 while providing syntactic subcategory information.

6 Discussion

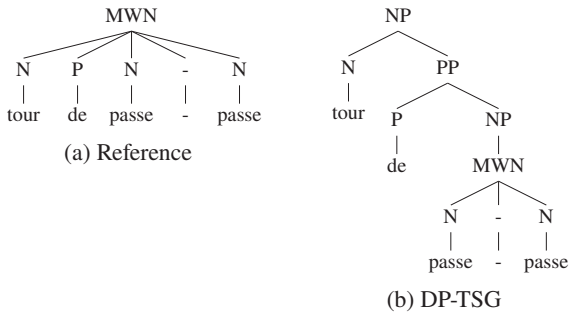
Automatic learning methods run the risk of producing uninterpretable models. However, the DP-TSG model learns useful generalizations over MWEs. A sample of the rules is given in Table 9. Some specific sequences like "[MWN [coup de N]]" are part of the grammar: such rules can indeed generate quite a few MWEs, e.g., *coup de pied* 'kick', *coup de coeur*, *coup de foudre* 'love at first sight', *coup de main* 'help', *coup d'état*, *coup de grâce* (note that only some of these MWEs are seen in the training data). For MWV, "V de N" as in *avoir de cesse* 'give no peace', *perdre de vue* [lose from sight] 'forget', *prendre de vitesse* [take from speed] 'outpace'), is learned. For prepositions, the grammar stores full subtrees of MWPs, but can also generalize the structure of very frequent sequences: "en N de" occurs in many multiword prepositions (e.g., *en compagnie de*, *en face de*, *en matière de*, *en terme de*, *en cours de*, *en faveur de*, *en raison de*, *en fonction de*). The TSG grammar thus provides a categorization of MWEs consistent with the Lexicon-Grammar. It also learns verbal phrases which contain discontinuous MWVs due to the insertion of an adverb or negation such as "[VN [MWV va] [MWADV d'ailleurs] [MWV bon train]]" [go indeed well], "[VN [MWV a] [ADV jamais] [MWV été question d']]" [has never been in question].

MWN
sociétés de N
prix de N
coup de N
N d'état
N de N
N à N

MWV
sous - V
faire N
V les moyens
V de N
V en N

MWP
de l'ordre de
y compris
au N de
en N de
ADV de

Table 9: Sample of the TSG rules learned.



(a) Reference (b) DP-TSG

Figure 2: Example of an MWE error for *tour de passe-passe* 'magic trick'. (dev set)

A significant fraction of errors for MWNs occur with adjectives that are not recognized as part of the MWE. For example, since *établissements privés* 'private corporation' is unseen in the training data, it is not found. Sometimes the parser did not recognize the whole structure of an MWE. Figure 2 shows an example where the parser only found a subpart of the MWN *tour de passe-passe* 'magic trick'.

Other DP-TSG errors are due to inconsistencies in the FTB annotation. For example, *sous prétexte que*

configuration details.