

Phrase Structure Grammar (Chomsky)	Dependency Grammar (Tesnière, Melcuk)
<i>Definite Clause Grammar (DCG) - Prolog</i> <i>Transformational-Generative Grammar (TGG)</i> <i>Head-Driven Phrase Structure Grammar (HPSG)</i> <i>Functional Unification Grammar (FUG),</i> <i>Lexical Functional Grammar (LFG)</i> <i>Tree Adjoining Grammar (TAG) - Arvid Joshi</i> <i>AGFL</i> ... <i>Functional Grammar (FG) - Simon Dik</i> <i>Systemic Functional Grammar (SFG) - Halliday</i>	<i>Constraint Grammar (CG)</i> <i>Functional Dependency Grammar (FDG)</i> <i>Topological Dependency Grammar (TDG)</i> <i>Extensible Dependency Grammar (XDG)</i> <i>Link Grammar (D-H symmetry)</i> <i>Dependency Grammar Annotator (DGA)</i>
constituent based structure	token based structure
explanatory-linguistic perspective generative tradition: parsers	descriptive-applicational perspective analytical tradition: taggers
(labeled) brackets	(labeled) arcs
rewriting rules	attachment rules
originally fixed word order languages <ul style="list-style-type: none"> English, French 	originally free word order languages <ul style="list-style-type: none"> Slavonic languages, Finnish, German
Problems: Discontinuity, free word order	Problems: Coordination, ellipsis
declarative programming	procedural programming
linguist-written: AGFL, HPSG, VISL-PSG, PATR, XTAG , ... machine-learned: PCFG e.g. Viterbi, Collins, Bikel	linguist-written: ENGCG, GERCG (Lingsoft), Machineese parsers (Conexor), VISL parsers machine-learned: MALT, Matsumoto, MSTParser,

	Constituent treebanks	Dependency treebanks
English	Penn I & II, Susanne Corpus , TOSCA Lancaster/IBM treebank (Spoken E.)	English Dependency Treebank and The PARC 700 Dependency Bank and CHILDES (Brown)
Arabic	Penn Arabic Treebank	PDT-Arabic (Smrž et.al. 2002)
Basque		Eus3LB
Bulgarian	BulTreeBank (Simov et.al. 2005) HPSG	-->
Catalan		Cat3LB

	Constituent treebanks	Dependency treebanks
Chinese	Penn Chinese Treebank	Sinica Treebank (Chen et.al. 2003)
Czech		PDT (Hajič et.al. 2001)
Danish	Arboretum (Bick 2003)	Arboretum (Bick 2003), Danish Dependency Treebank (Kromann 2003)
Dutch	Corpus Gesproken Nederlands	Alpino Treebank (van der Beek et.al. 2002)
Estonian	Arbores	
French	IBM Paris Treebank, Abeillé Treebank , L'Arboratoire	
German	NEGRA , TIGER Treebank (Brants et.al 2002), TueBa-D/S (spoken, +topology), TueBa-D/Z (written)	-->
Greek		Greek Dependency Treebank
Hebrew	Hebrew Treebank (a la Penn)	
Hindi		
Hungarian	(project: Hungarian treebank)	
Italian	VIT - Venice Italian Treebank	Turin University Treebank (Bosco et.al.)
Japanese	VERBMOBIL (Kawata and Bartels 2000)	ATR Dependency corpus (Lepage et.al.)
Korean	(project: Korean Treebank)	
Norwegian	(project: TREPIL Norwegian treebank)	
Polish	(HPSG test suite treebank)	
Portuguese	Floresta Sintá(c)tica (Afonso et.al. 2002)	-->
Slovene		Slovene Dependency Treebank (Džerosky et.al. 2006)
Spanish/ Catalan	UAM Treebank of Spanish	Cast3LB treebank (Palomar et.al. 2004)
Swedish		Talbanken95 (Nilsson et.al. 2005)
Turkish		METU-Sabancı (Ofłazer et.al. 2003)
Parallel	Sofie Treebank	

Parsers

	<i>CG</i> <i>(manual rules)</i>	<i>statistical DG</i> <i>(trained on treebanks)</i>	<i>PSG</i> <i>(manual rules)</i>	<i>PCFG</i> <i>(trained on treebanks)</i>
English	ENGCG (Lingsoft), FDG (Connexor) EngGram (VISL)	MALT (Nivre), Matsumoto, MSTParser (McDonald), Covington (Prolog, Chart Parser), adapted Collins (unlabeled, projective)	AlveyNLT (Unification based), AGFL (affix grammars)	Viterby PCFG (best tree), Bottom-up PCFG parser (forest), Collins, Bikel (head-driven)
Arabic			AGFL	Bikel
Bulgarian				
Chinese				Bikel
Czech		MALT, MSTParser		Collins
Danish	DanGram			
Dutch			AGFL	
Esperanto	EspGram			
Finnish	+			
French	FrAG			Collins
German	GERCG (Lingsoft), GerGram			Collins
Italian	ItaGram Turin DG			
Japanese		CaboCha (Matsumoto)		
Latin			AGFL	
Portuguese	PALAVRAS			
Russian			AGFL	
Slovene				
Spanish	PALAVRAS-HIS			Collins
Swedish	SWECEG	MALT		
Turkish				

References

- Afonso, S., E. Bick, R. Haber and D. Santos. Floresta Sintá(c)tica: A treebank of Portuguese. In *Proceedings of LREC'02*. pp. 1698-1703 . Paris: ELRA
- van der Beek, L. G. Bouma, R. Malouf, G. van Noord. 2002. The Alpino Dependency Treebank. In: *Computational Linguistics in the Netherlands CLIN 2001*. pp. 8-22. Rodopi
- Bick, Eckhard. 2005. Turning Constraint Grammar Data into Running Dependency Treebanks. In: Civit, Montserrat & Kübler, Sandra & Martí, Ma. Antònia (red.), *Proceedings of TLT 2005, Barcelona*. pp.19-2
- Bikel, Daniel M.. [Design of a Multi-lingual, Parallel-processing Statistical Parsing Engine](#), HLT 2002.
- C. Bosco, V. Lombardo. Dependency and relational structure in treebank annotation. In Proceedings of Workshop on Recent Advances in Dependency Grammar at COLING'04, Geneve, Switzerland, 2004, [ps-zip](#)
- Brants, S., S. Dipper, S. Hansen, W. Lezius, G. Smith. 2002. The TIGER Treebank. *Proc. of TLT1*, Sozopol
- Briscoe, E. & J. Carroll (1993) "Generalised probabilistic LR parsing of natural language (corpora) with unification-based grammars." *Computational Linguistics*, 19(1). 25-60.
- Charniak, Eugene. A Maximum-Entropy-Inspired Parser. NAACL 2000
- Covington, Michael A.. *A dependency parser for variable-word-order languages*. Technical Report AI-1990-01, Artificial Intelligence Programs, The University of Georgia Athens, Georgia 30602 USA, January 1990.
- Covington, M. A. (2001) A Fundamental Algorithm for Dependency Parsing. In *Proceedings of the 39th Annual ACM Southeast Conference*, pp. 95-102.
- Džerosky, S., T. Erjavec, N. Ledinek, P. Pajas, Z. Žabokrtsky, A. Žele. 2006. Towards a Slovene Dependency Treebank. In *Proc. of LREC'06*, Genoa
- Collins, Michael. 1996. A New Statistical Parser Based on Bigra Lexical Dependencies. ACL 1996
- Collins, Michael. 1999. Head-Driven Statistical Models for Natural Language Parsing (Ph.D. Thesis). University of Pennsylvania
- Daelemans, W. and Van den Bosch, A. (2005) *Memory-Based Language Processing*. Cambridge University Press.
- Michael A. Covington, [A free-word-order dependency parser in Prolog](#), 2003
- Hajič, J., B. Hladká, and P. Pajas. 2001. The Prague Dependency Treebank: Annotation Structure and Support. In *Proc. of the IRCS Workshop on Linguistic Databases*, pp. 105-114. University of Pennsylvania.
- Karlsson, Fred, Atro Vouitilainen, Jukka Heikkilä and A. Anttila. 1995. Constraint Grammar - A Language-Independent System for Parsing Unrestricted Text. Mouton de Gruyter: Berlin.
- Kawata, Y and J. Bartels. 2000. *Stylebook for the Japanese Treebank in VERBMOBIL*. Universität Tübingen: *Verbmobil-Report 240*.
- Keh-Jiann, Chu-Ren Huang, Feng-Yi Chen, Chi-Ching Luo, Ming-Chung Chang, Chao-Jan Chen, and Zhao-Ming Gao. 2003. Sinica Treebank: Design Criteria, Representational Issues and Implementation. In A. Abeille (ed.) *Treebanks Building and Using Parsed Corpora*. Dordrecht:Kluwer, pp231-248.
- Kromann, M. T. 2003. The Danish Dependency Treebank. In J. Nivre and E. Hinrichs (ed.) *Proceedings of TLT2003*. Växjö University Press, Sweden
- Kudo, Taku & Yuji Matsumoto, [Japanese Dependency Analysis using Cascaded Chunking](#), CoNLL 2002
- McDonald, Ryan, & Koby Crammer & Fernando Pereira, [Online Large-Margin Training of Dependency Parsers](#), 43rd Annual Meeting of the Association for Computational Linguistics, 2005.
- Nart, B. Atalay, Kemal Oflazr, Bilge Say. 2003. The Annotation Process in the Turkish Treebank. In *Proceedings of the EACL Workshop on Linguistically Interpreted Corpora - LINC 2003*. Budapest
- Nilsson, J, J. Hall and J. Nivre. 2005. MAMBA Meets TIGER: Reconstructing a Swedish Treebank from Antiquity. In *Proceedings NODALIDA 2005*. Joensuu
- Nivre, Joakim & Scholz, Mario. 2004. Deterministic Dependency Parsing of English Text. In Proceedings of COLING 2004, August 23-27, 2004, Geneva, Switzerland, pp. 64-70
- Oflazer, K., B. Say, D.Z. Hakkani-Tür, G. Tür. 2003. Building a Turkish Treebank. In A. Abeillé (ed.) *Building and Exploiting Syntactically-annotated Corpora*. Kluwer
- Palomar, M. et. al. 2004. Construcción de una base de datos de árboles sintáctico-semánticos para el catalán, euskera y español. In: *Proceedings of SEPLN XX*, pp 81-88. Barcelona: ISSN 1135-5948

- Simov, K., P. Osenova, A. Simov, M. Kouylekov. 2004. Design and Implementation of the Bulgarian HPSG-based Treebank. In E. Hinrichs and K. Simov (ed.), *Journal of Research on Language and Computation*, Vol. 2, No. 4 , pp. 495-522. Kluwer
- Smrž, Otakar, Jan Šnidauf, and Petr Zemánek. 2002. Prague Dependency Treebank for Arabic: Multi-Level Annotation of Arabic corpus. In *Proceedings of the International Symposium on Processing of Arabic*, pages 147-155, Manouba, Tunisia, April 2002.
- Yamada, Hiroyasu and Yuji Matsumoto, [Statistical dependency analysis with Support Vector Machines](#), In Proc. 8th International Workshop on Parsing Technologies (IWPT), pp.195-206, April 2003.

Links

Constraint Grammar: http://beta.visl.sdu.dk/constraint_grammar.html

AGFL: <http://www.cs.ru.nl/agfl/>

Link Grammar: <http://www.link.cs.cmu.edu/link/> (English)

Alvey Natual Language Tools: <http://www.cl.cam.ac.uk/Research/NL/anlt.html> (English)

NLTK Tutorial (Edward Loper): Probabilistic Parsing: <http://nltk.sourceforge.net/tutorial/pcfg/index.html> (English)

TiMBL, Tilburg memory based learner: <http://ilk.uvt.nl/software.html> (tagging and dependency)

Constraint-based grammar formalisms: <http://cslu.cse.ogi.edu/HLTsurvey/ch3node5.html#SECTION33>