

Learning method	Application domain	Reading
Introduction		
Non-probabilistic knowledge		
Learnability theory		
Gold's Theorems		
Probably Approx. Correct Learnability		
Formal background		
Probability theory		
Information theory	Parameter setting (V2)	Brill & Kapur 1997
	Selectional restrictions	Resnik 1996; Nicol 2007
	Allophony	Peperkamp, Le Calvez, Nadal & Dupoux to appear
Clustering		
Continuous	Phoneme boundaries	Kuhl 2000; Maye & Gerken 2001
Discrete	Syntactic category	Mintz, Newport & Bever 2002
Intro to hidden structure	learning	
Generative models		
Expectation-Maximization		
Sequential (visible) structure		
Markov Models	Part-of-speech tagging	Kim, Srinivas, Trueswell 2002
Decision Trees		
Constraints		
Maximum-Entropy Models	Phonotactics	Hayes & Wilson to appear
	OT grammar learning	Goldwater & Johnson 2003
	Word Learning	Tenenbaum & Xu 2000, to appear
	Concept formation	Tenenbaum & Griffiths 2001
	Syntax learning	Perfors, Tenenbaum & Regier 2006
	Segmentation	Brent 1999
	Causal reasoning	
	Speech	
Hidden structure		
Bayesian Learning		
Bayesian Networks		
Hidden Markov Models		
Grammar learning		
Finite-State Automata	Phonological rules	Gildea & Jurafsky 1996
Probabilistic Context-Free Grammars	Syntax/Syllable Structure	Klein & Manning 2002, 2004/Goldwater & Johnson 2005
OT grammars & lexicons	Phonology	Jarosz 2006

References

- Brent, Michael. 1999. An efficient, probabilistically sound algorithm for segmentation and word discovery. *Machine Learning*, 34(1):71-106.
- Brill, Eric and Shyam Kapur. 1993. An information-theoretic solution to parameter setting. Technical Report IRCS-93-07, Institute for Research in Cognitive Science, University of Pennsylvania.
- Gildea, Daniel and Daniel Jurafsky. 1996. Learning bias and phonological-rule induction. *Computational Linguistics*, 22(4):497-530.
- Goldwater, Sharon and Mark Johnson. 2003. Learning OT Constraint Rankings Using a Maximum Entropy Model, In *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Eds: Jennifer Spenser, Anders Eriksson, and Östen Dahl. pp.111-120.
- Goldwater, Sharon and Mark Johnson (2005) Representational Bias in Unsupervised Learning of Syllable Structure, *Proceedings of the Workshop on Psychocomputational Models of Human Language Acquisition*, ACL 2005.
- Hayes, Bruce and Colin Wilson. To appear. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*.
- Jarosz, Gaja. 2006. Richness of the Base and Probabilistic Unsupervised Learning in Optimality Theory. Association for Computational Linguistics: *Proceedings of the Eighth Meeting of the ACL Special Interest Group in Computational Phonology*.
- Kim, Albert E., Bangalore Srinivas, and John C. Trueswell. 2002 A computational model of the grammatical aspects of word recognition as supertagging. In Suzanne Stevenson and Paola Merlo, editors, *The Lexical Basis of Sentence Processing: Formal, Computational and Experimental Issues*, pages 109-136. John Benjamins.
- Klein, Dan and Chris Manning. 2002. A generative constituent-context model for improved grammar induction. In Proc. 40th Annual Meeting Association for Computational Linguistics, pp. 128-135.
- Klein, Dan and Chris Manning. 2004. Corpus-based induction of syntactic structure: Models of dependency and constituency. In Proc. 42nd Annual Meeting Association for Computational Linguistics, pp. 479-486
- Kuhl, Patricia. 2000. A New View of Language Acquisition. *Proceedings of the National Academy of Sciences* 97:11850-11857.
- Maye, Jessica and LouAnn Gerken. 2001. Learning Phonemes: How Far Can the Input Take Us? *Proceedings of the BU Conference on Language Development*. 480-490.
- Mintz, T.H., Newport, E.L., & Bever, T.G. 2002. The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, 26:393-425.
- Peperkamp, S., Le Calvez, R., Nadal, J.P. and Dupoux, E. 2006. The acquisition of allophonic rules: statistical learning with linguistic constraints. *Cognition* 101(3):B31-B41.
- Perfors, A., Tenenbaum, J., and Regier, T. 2006. Poverty of the Stimulus? A rational approach. *28th Annual Conference of the Cognitive Science Society*.
- Resnik, Philip. 1996. Selectional Constraints: An Information-Theoretic Model and its Computational Realization, *Cognition*, 61:127-159.
- Tenenbaum, J.B. and T. L. Griffiths 2001. Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24:629-641.
- Tenenbaum, J.B. and F. Xu. 2000, Word learning as Bayesian inference. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*.
- F. Xu and J. B. Tenenbaum,. In press, Word learning as Bayesian inference. *Psychological Review*.