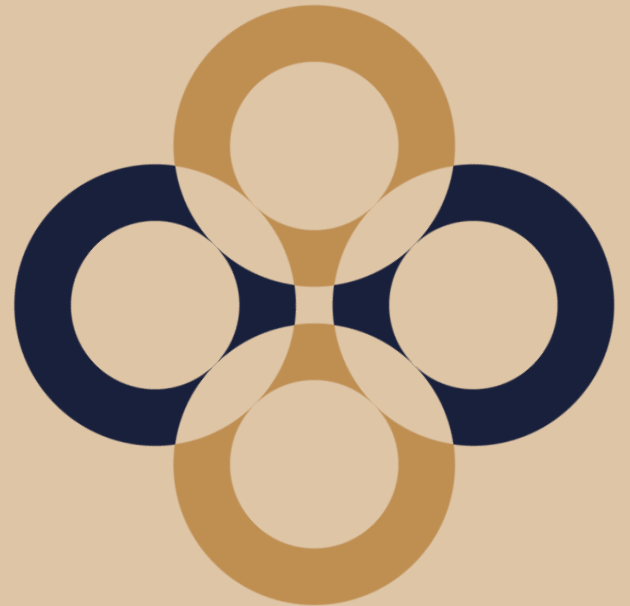


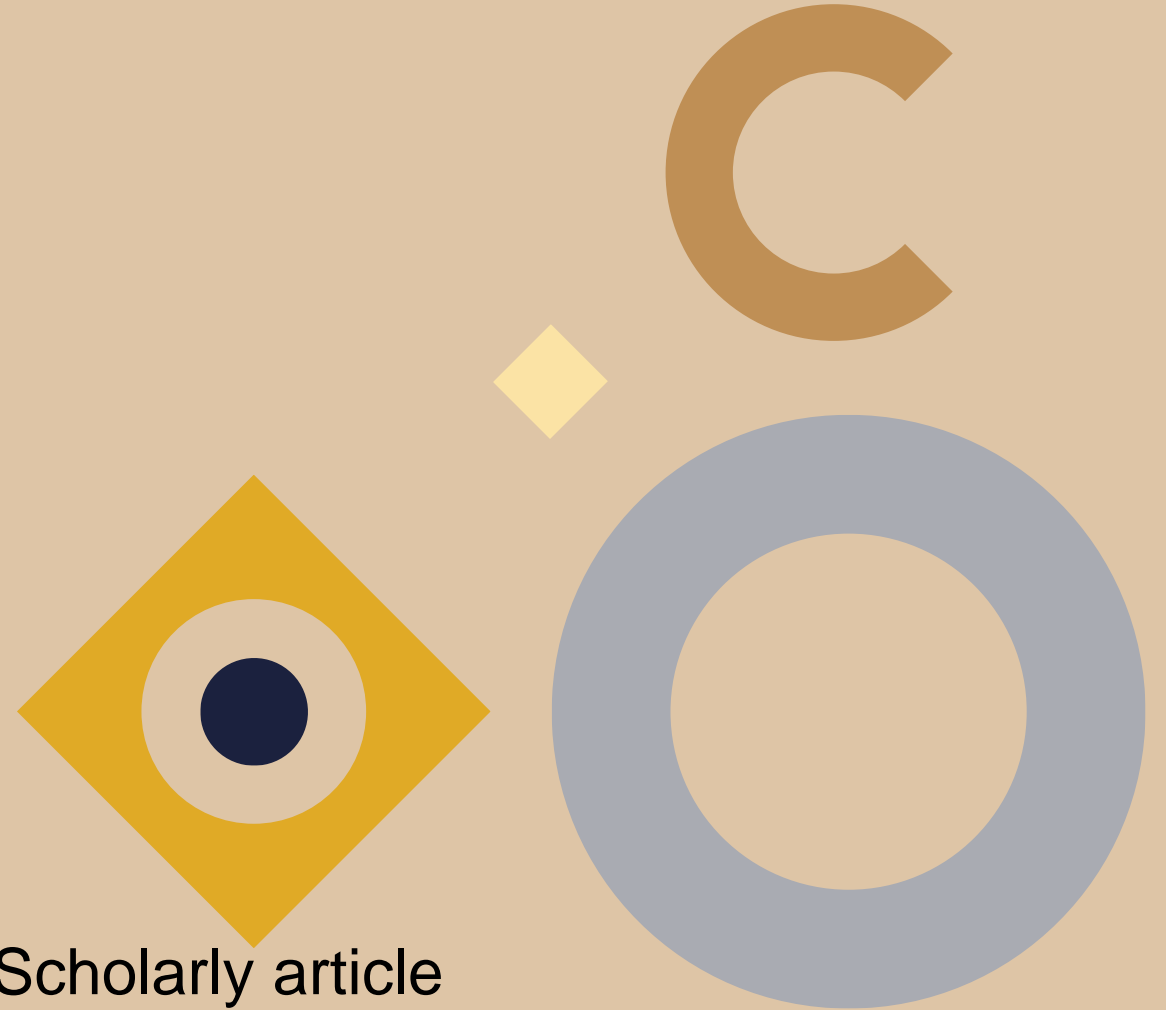
# Thesis Writer's Guide I.

Sources



# Content

- Topic selection
- Sources
- Where to look for what?
- SuperSearch, Topic sentences
- External sources, Source evaluation, Scholarly article



# Topic selection

- Own topic (interest, personal issue, current theme)
- Topic offered by department/supervisor/consultant
  
- General hints
  - Time management
  - Check sources in advance
  - Check former theses in [Repository](#)
    - formal requirements
    - theses' bibliography can serve as a starting point
  - Critical reading, questions
  - Narrow your topic (focus on it)



## Academic, scholarly literature

- Books
- Peer-reviewed articles



Source: <https://paperpile.com/g/find-credible-sources/>

## Grey literature

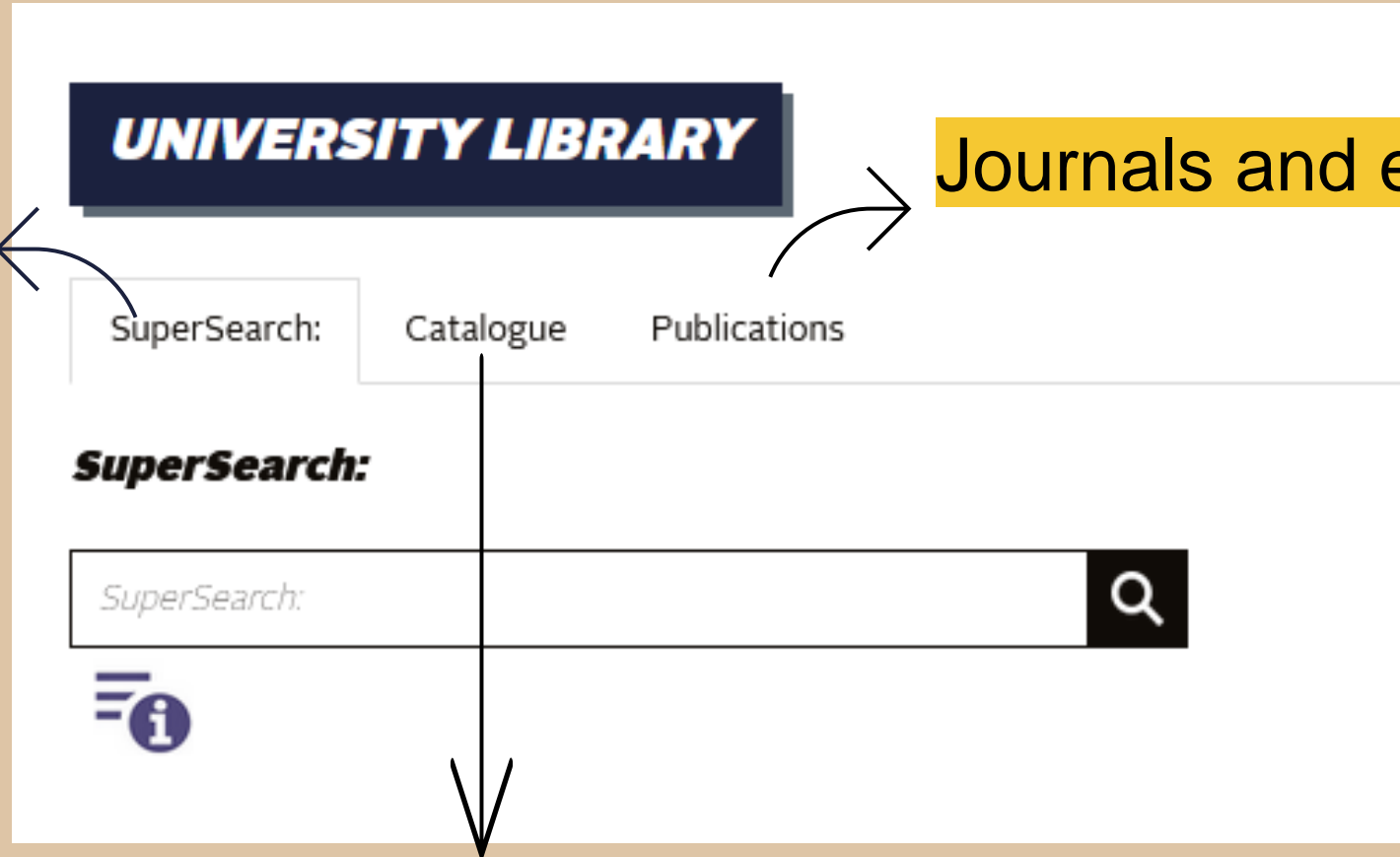
- non-commercially published
- non peer-reviewed
- e.g. working papers, company financials, theses, dissertations

## Data, figures

- Governmental, international or commercial organizations, statistical databases
- Privately collected (surveys)
- [SPSS](#)

# Where to look for what? - Library sources

For topic search



Journals and ebooks

+Databases

+Repositories

Mainly printed books + e-links

### My dashboard

- Overview
- Projects
- Saved
- Searches
- Viewed

### Research tools

- General search
- Publications
- Concept map
- Supplemental sources

# SuperSearch - (Live demonstration)

## Building a search query

- Create a topic sentence
- Pick up some keywords, find synonyms
- Use search techniques: logical operators (AND, OR, NOT), „phrase search”, truncation (\*)

## Search articles, books, journals & more

Search for topic using keywords

Peer reviewed

- Fine tuning with filters, field search
- Save results into Dashboard/export to Zotero

# Examples of topic sentences

*Role of banks in financial crisis*

*Corporate Social Responsibility Practices in the Energy Industry*

*Health tourism enterprises and adaptation for sustainable development*

# External sources – Source evaluation

- Interlibrary loan, Supplemental sources, Concept Map
- „To Google or not to Google?”, Wikipedia
- Google Scholar

## Source evaluation (CRAAP-test)

<b>C</b>	<p><b>Currency: <i>The timeliness of the information.</i></b></p> <ul style="list-style-type: none"> <li>• When was the information published or posted? Revised or updated?</li> <li>• Does your topic require current information, or will older sources work as well?</li> </ul>
<b>R</b>	<p><b>Relevance: <i>The importance of the information for your needs.</i></b></p> <ul style="list-style-type: none"> <li>• Does the information relate to your topic or answer your question?</li> <li>• Who is the intended audience? / an appropriate level?</li> </ul>
<b>A</b>	<p><b>Authority: <i>The source of the information.</i></b></p> <ul style="list-style-type: none"> <li>• Who is the author/publisher/source/sponsor?</li> <li>• What are the author's credentials or organizational affiliations?</li> <li>• Is the author qualified to write on the topic? / contact information?</li> </ul>
<b>A</b>	<p><b>Accuracy: <i>The reliability, truthfulness and correctness of the content.</i></b></p> <ul style="list-style-type: none"> <li>• Where does the information come from? / supported by evidence?</li> <li>• Has the information been reviewed or refereed?</li> <li>• Does the language or tone seem unbiased and free of emotion?</li> </ul>
<b>P</b>	<p><b>Purpose: <i>The reason the information exists.</i></b></p> <ul style="list-style-type: none"> <li>• What is the purpose of the information? Is it to inform, teach, sell, entertain or persuade?</li> <li>• Does the point of view appear objective and impartial?</li> <li>• Are there political, religious, institutional or personal biases?</li> </ul>



Source: <https://www.emaze.com/@AIFFRWC/C.R.A.A.P.-Test-for-Evaluating-Websites>



# Signs of a scholarly article

Title

A Cognitive Model for the Representation and Acquisition of Verb Selectional Preferences

Author Details

Afra Alshahi  
Department of Computer Science  
University of Toronto  
afra@cs.toronto.edu

Suzanne Stevenson  
Department of Computer Science  
University of Toronto  
suzanne@cs.toronto.edu

Abstract

**Abstract**

We present a cognitive model of inducing verb selectional preferences from individual verb usages. The selectional preferences for each verb argument are represented as a probability distribution over the set of semantic properties that the argument can possess—a *semantic profile*. The semantic profiles yield verb-specific conceptualizations of the arguments associated with a syntactic position. The proposed model can learn appropriate verb profiles from a small set of noisy training data, and can use them in simulating human plausibility judgments and analyzing implicit object alternation.

Introduction

**1 Introduction**

Verbs have preferences for the semantic properties of the arguments filling a particular role. For example, the verb *eat* expects that the object receiving its theme role will have the property of being edible, among others. Learning verb selectional preferences is an important aspect of human language acquisition, and the acquired preferences have been shown to guide children's expectations about missing or upcoming arguments in language comprehension (Nation et al., 2003).

Resnik (1996) introduced a statistical approach to learning and use of verb selectional preferences. In this framework, a semantic class hierarchy for words is used, together with statistical tools, to induce a verb's selectional preferences for a particular argument position in the form of a distribution

Publication

Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition, pages 41–48, Prague, Czech Republic, June 2007. ©2007 Association for Computational Linguistics

over all the classes that can occur in that position. Resnik's model was proposed as a model of human learning of selectional preferences that made minimal representational assumptions; it showed how such preferences could be acquired from usage data and an existing conceptual hierarchy. However, his and later computational models (see Section 2) have properties that do not match with certain cognitive plausibility criteria for a child language acquisition model. All these models use the training data in "batch mode", and most of them use information theoretic measures that rely on total counts from a corpus. Therefore, it is not clear how the representation of selectional preferences could be updated incrementally in those models as the person receives more data. Moreover, the assumption that children have access to a full hierarchical representation of semantic classes may be too strict. We propose an alternative view in this paper which is more plausible in the context of child language acquisition.

In previous work (Alshahi and Stevenson, 2005), we have proposed a usage-based computational model of early verb learning that uses Bayesian clustering and prediction to model language acquisition and use. Individual verb usages are incrementally grouped to form emergent classes of linguistic constructions that share semantic and syntactic properties. We have shown that our Bayesian model can incrementally acquire a general conception of the semantic roles of predicates based only on exposure to individual verb usages (Alshahi and Stevenson, 2007). The model forms probabilistic associations between the semantic properties of arguments, their syntactic positions, and the semantic primitives

Alternating verbs		Non-alternating verbs	
write	0.61	long	0.76
sing	0.67	wear	0.71
drink	0.67	try	0.75
eat	0.74	catch	0.76
play	0.74	allow	0.77
pass	0.76	make	0.78
watch	0.77	let	0.78
push	0.78	open	0.81
send	0.80	take	0.83
push	0.80	see	0.87
call	0.80	like	0.87
pull	0.80	get	0.87
explode	0.81	find	0.87
read	0.82	give	0.88
learn	0.87	bring	0.89
		want	0.89
		put	0.90
Mean	0.76	Mean	0.81

Figure 6: Similarity with the base profile for Alternating and Non-alternating verbs.

than verbs with stronger preferences. We use the cosine measure to estimate the similarity between two profiles  $p$  and  $q$ :

$$\text{cosine}(p, q) = \frac{p \cdot q}{|p| \times |q|} \quad (9)$$

The similarity values for the Alternating and Non-alternating verbs are shown in Figure 6. The larger values represent more similarity with the base profile, which means a weaker selectional preference. The means for the Alternating and Non-alternating verbs were respectively 0.76 and 0.81, which confirm the hypothesis that verbs participating in implicit object alternations select more strongly for the direct objects than verbs that do not. However, like Resnik (1996), we find that it is not possible to set a threshold that will distinguish the two sets of verbs.

## 5 Conclusions

We have proposed a cognitively plausible model for learning selectional preferences from instances of verb usage. The model represents verb selectional preferences as a semantic profile, which is a probability distribution over the semantic properties that an argument can take. One of the strengths of our model is the incremental nature of its learning mechanism, in contrast to other approaches which learn selectional preferences in batch mode. Here we have only reported the results for the final stage of learning, but the model allows us to monitor the semantic

profiles during the course of learning, and compare it with child data for different age groups, as we do with semantic roles (Alshahi and Stevenson, 2007). We have shown that the model can predict appropriate semantic profiles for a variety of verbs, and use these profiles to simulate human judgments of verb-argument plausibility, using a small and highly noisy set of training data. The model can also use the profiles to measure verb-argument compatibility, which was used in analyzing the implicit object alternation.

## References

Alshahi, S. and Light, M. (1999). Hiding a semantic hierarchy in a Markov model. In *Proc. of the ACL Workshop on Unsupervised Learning in Natural Language Processing*.

Alshahi, A. and Stevenson, S. (2005). A probabilistic model of early argument structure acquisition. In *Proc. of the CogSci 2005*.

Alshahi, A. and Stevenson, S. (2007). A computational usage-based model for learning general properties of semantic roles. In *Proc. of the EuroCogSci 2007*.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, 98(3): 409–423.

Brockmann, C. and Lapata, M. (2003). Evaluating and combining approaches to selectional preference acquisition. In *Proc. of the EACL 2003*.

Caramazza, M. and Johnson, M. (2000). Explaining away ambiguity: Learning verb selectional preference with Bayesian networks. In *Proc. of the COLING 2000*.

Clark, S. and West, D. (2002). Class-based probability estimation using a semantic hierarchy. *Computational Linguistics*, 28(2):187–210.

Collins, M. (1999). *Head-Driven Statistical Models for Natural Language Parsing*. PhD thesis, University of Pennsylvania.

Hilfiker, V. M., Snow, L., and Cappelis, L. (1989). Lexical expectations in parsing complement-verb sentences. *Journal of Memory and Language*, 28:608–630.

Levin, B. (1995). *English verb classes and alternations: A preliminary investigation*. The University of Chicago Press.

Li, H. and Abe, N. (1998). Generalizing noun frames using a theorem and the MDL principle. *Computational Linguistics*, 24(2):107–122.

Light, M. and Grillo, W. (2002). Statistical models for the induction and use of selectional preferences. *Cognitive Science*, 26(3):209–231.

MacWhinney, B. (1995). *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum.

Miller, G. (1990). WordNet: An on-line lexical database. *International Journal of Lexicography*, 13(3).

Nation, K., Marshall, C. M., and Alstrom, G. Y. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-evoked eye movements. *J. of Experimental Child Psych.*, 86:314–329.

Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realization. *Cognition*, 61:127–159.

Charts & Equations

References

Body Text

Conclusions



**Thank you for  
your attention!**