Comparing Computational Models of Selectional Preferences —
Second-order Co-Occurrence vs. Latent Semantic Clusters

Sabine Schulte im Walde
schulte@ims.uni-stuttgart.de
SFB 732 "Incremental Specification in Context"
Project D4 "Modular Lexicalisation of Probabilistic Context-Free Grammars"
Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart

Second-order Co-Occurrence Selectional Preferences

Idea
• Selectional preferences with respect to a predicate’s complement are defined by the properties of the complement realisations.
• Example: typical direct object of drink is fluid, might be hot or cold, can be bought, might be bottled, etc.
• Second-order co-occurrence: a predicate’s restrictions to the semantic realisation of its complements are expressed through the properties of the complements.

Scoring
• Selectional preference description:
  \[ \text{score}(v,n,prop) = \sum \log(\text{freq}(f, i, c)) \log(\text{freq}(x, v, i, c)) \]
  Variations of frequency: log(freq) and prob.
• Second-order preference fit of a noun by standard distributional measures: compares noun’s contribution to overall preference.
  \[ \cosine, skew divergence, tau, jaccard \]

Evaluation
Data:
• Human judgements on 90 German subjects, direct objects, pp objects (across 30 verbs).
• Taken from Brockmann & Lapata (2003).
• Correlation of system scores with human judgements, by linear regression.
• Brockmann & Lapata (BL) normalised system scores by log10.
Baselines and Upper Bound:
• Correlation of joint corpus-based predicate-noun frequencies with judgments.
• Two baselines: raw frequencies and frequencies transformed by log10.
• Upper bound: inter-subject agreement on selectional preference judgements.

Example judgements:

Example: typical direct object of drink is fluid, might be hot or cold, can be bought, might be bottled, etc.

Selectional Restrictions
• Predicates impose selectional restrictions on their complements.
• Example: (Chomsky, 1957): Colorless green ideas sleep furiously.
• Syntactically well-formed but semantically meaningless.
• Realisation of complements with respect to thematic role.
• Examples: Elsa baked a chocolate cake, *Elsa baked a stone.

Selectional Preferences
• Degree of acceptability.
• Probabilistic models.

Computational Approaches
Cluster-based vs. WordNet-based vs. distributional.

LSC (Latent Semantic Clusters)
• Two-dimensional clustering model.
• Self-clustering approach.
• Training by EM algorithm (Baum, 1972).
• Probability model for verb-noun pairs:
  \[ p(c,v,n) = \prod p(v,c) p(n,c) \]

PAC (Predicate-Argument Clustering)
• Extension of LSC model by selectional preferences.
• Incorporates Minimum Description Length (MDL) cuts through WordNet hierarchy.
• Probability model for verb-noun pairs via frame types:
  \[ p(v,f,a_1,...,a_n) = \prod p(v,c) p(l,v,c) p(f,l,c) \prod_{i} [p(r_1,c,f,a_i) p(a_i,l)] \]