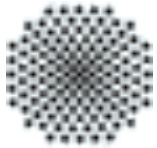
Characterising Response Types and Revealing Noun Ambiguity in German Association Norms



Sabine Schulte im Walde

(University of Stuttgart)

Joint work with Alissa Melinger and Andrea Weber

(University of Dundee and Saarland University)

Universitat Pompeu Fabra

April 20, 2007

Overview

1. Motivation, research questions, and linguistic intuitions: processing ambiguity in association experiments
2. Noun association data collection
3. Statistical analysis of association response types
4. Cluster analysis of noun senses
5. Conclusions

Psycholinguistic experiments ⇔ computational methods

Processing Ambiguity

- Language is rife with ambiguity.
- Humans are not greatly disturbed.
  Experimental work investigates how we cope.
- When processing ambiguous words, multiple meanings are initially activated, even in highly constraining contexts. (Tanenhaus et al. 1979)
  - They all *rose*.
  - They bought a *rose*.

  Both sentences prime semantically related words such as *flower* and *stand.*

Differences by Modality

- Comparison of picture vs. word semantic processing.
- Differences in semantic information (→ associations), elicited in the two presentation modes.
- The associations should reflect
  - ambiguity,
    if multiple meanings of a word are activated;
  - degrees of activation of the various meanings.

Association Norms

- Target stimuli are presented to experiment participants.
- The participants provide associate responses, i.e., words that are called to mind by the target words.

Example: *schneien* → *kalt, Schnee, Winter, weiß*

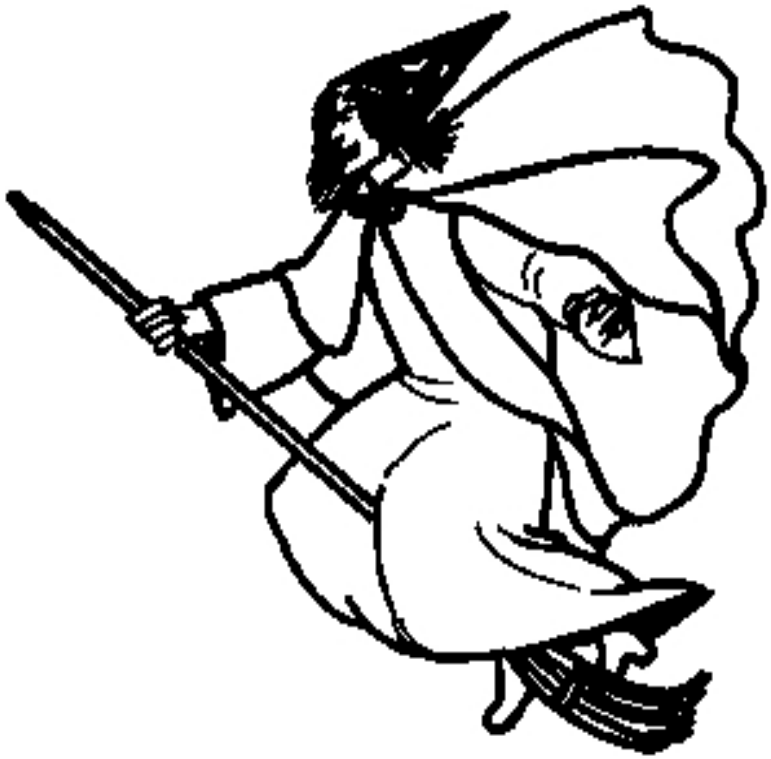
- This work is based on a collection of association norms for German nouns, using lexical and pictorial stimuli.

Linguistic Intuitions

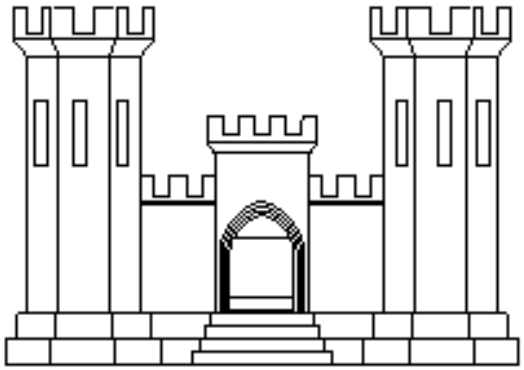
Associate responses elicited by written words are different from associate resonses elicited by pictures:

- Images might increase the salience of physical attributes of objects.
- Images might show non-prototypical characteristics of objects that would not be evoked by words.
- When word forms have different shades of meaning, responses evoked by lexical stimuli might index any of the words‘ meanings while responses evoked by pictorial representations might be more biased towards the depicted sense.

Which *Witch?*



Which *Schloss?*



Noun Association Analyses

1. Are there systematic differences in associate response types when target objects are presented in written form compared to when the written form is accompanied by a pictorial representation?

   → analysis and comparison of response types

2. Can we identify multiple senses of the nouns and discriminate between noun senses based on the associate responses?

   → cluster analysis of target-response pairs

Experiment Material

- 409 German concrete nouns
- Depictable objects
- Variety of semantic categories:
  - » plants: *Rose* `rose´, *Baum* `tree´, *Zweig* `branch´
  - » professions: *Doktor* `doctor´, *Bäcker* `baker´
  - » instruments: *Klavier* `piano´, *Trommel* `drums´
  - » body parts: *Auge* `eye´, *Kopf* `head´, *Fuß* `foot´ ...
- Homophones: ca. 10% of the nouns
- Variety of frequency ranges according to CELEX

Association Elicitation

- 409 stimuli divided into 3 questionnaires
- Each set presented in two formats:
  with and without pictures
- 300 native German participants;
  50 participants for each questionnaire
- Maximum of three associates per stimulus
- No time limit

Modality: *picture + word (PW)*

Witch

magic

wizard

broom

Modality: *word only (W)*

Witch

magic

wizard

wicked

Response Type Distribution

Schloss `lock´ (depicted), `castle´

Association		POS	PW	W
Schlüssel	‘key’	N	38	13
Tür	‘door’	N	10	5
Prinzessin	‘Princess’	N	0	8
Burg	‘castle’	N	0	8
sicher	‘safe’	ADJ	7	0
Fahrrad	‘bike’	N	7	0
schließen	‘close’	V	6	1
Keller	‘cellar’	N	7	0
König	‘king’	N	0	7
Turm	‘tower’	N	0	6
Sicherheit	‘safety’	N	5	1

Number of Response Tokens

- Prediction: The overall number of response tokens is unlikely to differ for the two presentation modes, since participants are limited to three associate resonses per target stimulus in both presentation modes:
  Token(PW) ~ Token(W)
- Analysis: *Token(PW) > Token(W)* - 59%
  Token(PW) < Token(W) - 32%
  Token(PW) = Token(W) - 8%
  significant difference
- Conclusion:
  Pictures facilitate the production of associations.

Number of Response Types

- Prediction: The overall number of resonse types should differ. In the PW condition we expect a bias towards the depicted noun sense, resulting in a smaller number of response types than in the W condition:
  Type(PW) < Type(W)
- Analysis: *Type(PW) < Type(W)* - 56%
  Type(PW) > Type(W) - 37%
  Type(PW) = Type(W) - 7%
  significant difference
- Conclusion:
  Pictures encourage less response types.

Idiosyncratic Response Types

- Prediction: The PW condition should produce less idiosyncratic response types than the W condition, because pictures reinforce associations that are depicted or at least related to the picture:
  Idio(PW) < Idio(W)
- Analysis: *Idio(PW) < Idio(W)* - 53%
  Idio(PW) > Idio(W) - 43%
  Idio(PW) = Idio(W) - 4%
  significant difference
- Conclusion:
  Pictures enhance less response diversity.

Part-of Response Types

- Prediction: The PW condition should receive more associations that show a part-of relation to the target stimulus than the W condition, because characteristics of the pictures can highlight specific parts of the whole:
  Part(PW) > Part(W)
- Analysis: *Part(PW) > Part(W)* - 29%
  Part(PW) < Part(W) - 35%
  Part(PW) = Part(W) - 36%
  no significant difference
- Conclusion:
  Pictures do not enhance a part-whole relationship.

Type Agreement

- Prediction: The number of response types on which the PW and the W conditions agree is expected to differ with respect to the target noun. For highly ambiguous target nouns we expect low type agreement because there are more senses which can be addressed.
- This prediction does not refer to a PW-W distinction, but instead uses the PW-W distinction to approach the issue of noun senses.
- Analysis: The targets are sorted by proportion of type agreement, and the top and bottom 20 are compared. A random distribution predicts 2 homophones per set.

Calculating Type Agreement

noun_PW: Ball	*noun_W: Ball*
sports	round
round	toy
soccer	dance

Total number shared responses = 1 (round)
Total number of response types = 5
Proportion agreement = .2

Type Agreement Lists

10 **top** Targets			10 **bottom** Targets		
Target Noun		Amb.	Target Noun		Amb.
Bank	`bank, bench´	yes	*Keks*	`biscuit´	no
Becken	`basin, cymbal, ...´	yes	*Kerze*	`candle´	no
Blatt	`leaf, paper, ...´	yes	*Mond*	`moon´	no
Boje	`buoy´	no	*Pinsel*	`brush´	no
Dackel	`dachshund´	no	*Salat*	`salad´	no
Erde	`earth, soil´	yes	*Schaf*	`sheep´	no
Filter	`filter´	no	*Schal*	`scarf´	no
Fontäne	`fountain´	no	*Schlüssel*	`key´	no
Geige	`violin´	no	*Teddy*	`teddy´	no
Karton	`carton´	no	*Wespe*	`wasp´	no

9 ambiguous nouns in the top 20 list vs. 1 in the bottom 20 list.
→ significant difference

Summary of Statistical Analyses

- The associate responses for concrete German nouns differ significantly depending on the presentation mode.
- Alternative meanings seem more active with written vs. depicted stimuli.
- Nevertheless, alternative meanings are also available when processing pictorial stimuli with ambiguous nouns.
- Not all our intuitive predictions were born out, e.g., the part-of relation between targets and responses.

Analysis of Noun Senses

1. Soft clustering of target-response pairs:
   similar nouns (i.e., shared associates) in common clusters

2. Predict the ambiguity of nouns and their senses

3. Evaluate the clusters against the *Duden* dictionary;
   interannotator agreement

Latent Semantic Clusters (LSC)

- Expectation-Maximization (EM) algorithm (Baum, 1972)
- Unsupervised training on unannotated data
- Generalisation over hidden data
- Two-dimensional clusters:
  - » cluster probability
  - » *dim1:* probabilities for target nouns in clusters
  - » *dim2:* probabilities for associations in clusters
- Previous work (Rooth, 1998; Rooth et al., 1999): selectional dependencies in grammatical relationships

Application of LSC to Noun Data

- Basis: joint frequencies of target nouns and associations
- Target description: *noun_PW* vs. *noun_W*
- Number of clusters: 100 / 200
- Implementation by Helmut Schmid

LSC Cluster Example

cluster, p(c) = 0.01295 (range: 0.00530 - 0.02674)

Rüstung_W	`armour´	0.097
Schwert_W	`sword´	0.097
Burg_W	`castle´	0.096
Rüstung_PW	`armour´	0.096
Dolch_PW	`dagger´	0.095
Schwert_PW	`sword´	0.093
Burg_PW	`castle´	0.091
Dolch_W	`dagger´	0.089
Ritter_PW	`knight´	0.073
Ritter_W	`knight´	0.068
Schloss_W	`castle´	0.040
Turm_PW	`tower´	0.014

Ritter	`knight´	0.158
Mittelalter	`medieval times´	0.067
Rüstung	`armour´	0.041
Burg	`castle´	0.037
Kampf	`fight´	0.036
kämpfen	`fight´	0.027
Schwert	`sword´	0.026
Waffe	`weapon´	0.022
Schloss	`castle´	0.020
scharf	`sharp´	0.019
Blut	`blood´	0.018
spitz	`sharp´	0.018

Noun Ambiguity and Noun Senses

- Which associations are highly probable for a cluster?
  → semantic content of the cluster
- Which target nouns are highly probable for a cluster and ist semantic content, i.e. the associations?
  → noun senses
- Which target nouns are in the same cluster and therefore refer to a common sense/aspect of the nouns?
  → PW vs. W

Basis: 200-cluster analysis, probability cut-off: 1%

Prediction of Noun Ambiguity

- Ambiguity of target noun: number of clusters
- Example:
  - » *Becken* `basin, cymbal, pelvis´ (among others):
    member of 8 clusters
  - » *Bäcker* `baker´:
    member of 1 cluster
- 735 noun senses in 200 clusters

Discrimination of Noun Senses

- Associations in cluster discriminate noun sense
- Example: *Becken* `basin, cymbal, pelvis´ (among others)
  - » *Wasser* `water´, *Garten* `garden´, *Feuerwehr* `fire brigade, *gießen* `water´, *nass* `wet´ → basin
  - » *Musik* `music´, *laut* `loud´, *Instrument* `instrument´, *Orchester* `orchestra´, *Jazz* → music
  - » *Hand* `hand´, *Bein* `leg´, *Ellenbogen* `elbow´, *Körper* `body´, *Muskel* `muscle´ → body
- Polysemy vs. word facets,
  e.g. *Filter* `filter´: coffee, cigarette, car

Noun Similarity

- Semantic similarity: common cluster membership
- Example: *Becken* `basin, cymbal, pelvis´ (among others)
  - » *Eimer* `bucket´, *Fontäne* `fountain´, *Brunnen* `fountain, well´, *Weiher* `pond´, *Vase* `vase´ → basin
  - » *Tuba* `tuba´, *Trompete* `trumpet´, *Saxophon* `sax´, *Trommel* `drum´ → music
  - » *Arm* `arm´, *Knochen* `bone´ → body

Discrimination of *PW vs. W* Noun Senses

- Associations in cluster discriminate noun sense & semantic similarity: common cluster membership
- Same noun sense in PW and W condition:
  word_PW and *word_W* appear in the same cluster
- Noun sense in either PW or W condition:
  only *word_PW* or *word_W* appears in a cluster
- Example: *Becken* `basin, cymbal, pelvis´ (among others)
  - » *Becken_W* only → basin
  - » *Becken_PW* and *Becken_W* → music (depicted!)
  - » *Becken_W* only → body

Evaluation of Noun Clusters

- Goal: evaluation of predictability
- Task: annotate noun senses with *Duden* dictionary
- 2 annotators
- Selection from 409 target nouns:
  - » manual selection of 20 homophones
  - » 20 top nouns from type agreement list
  - » 20 bottom nouns from type agreement list

  → 51 target nouns
- 100 and 200 cluster analysis, cut-off: 1%

Duden Definitions (examples)

Becken

1. Wasch-/Toilettenbecken `washbasin´
2. Schwimmbecken `swimming pool´
3. (Geologie) Senke, Mulde `(geol.) basin´
4. Teil des Körpers `body part´
5. Musikinstrument `musical instrument´

Schloss

1. Vorrichtung zum Verschließen `device for closing´
2. Wohngebäude von Fürsten und Adeligen
   `residential building for princes and noblemen´

Sense Annotation (example)

***** ***Becken*** *****

KLASSE: Wasser, See, Fluss, schwimmen, Jeans

MEMBER: Boot, Biber, Fisch, Hose, Kanu, Paddel, Brunnen, Eimer, Schwan, Becken, Ente, Boje, Brücke, Vase, Fähre, Karaffe, Fluss, Segelboot, Gondel

Duden Sense: 2

KLASSE: laut, Venedig, Musik, schlagen, Schlagzeug

MEMBER: Trommel, Gondel, Becken, Gong

Duden Sense: 5

Sense Annotation (example)

***** ***Becken*** *****

CLASS: water, sea, river, swim, jeans

MEMBER: boat, beaver, fish, trousers, canoe, paddle, fountain, bucket, swan, basin, duck, buoy, bridge, vase, ferry, carafe, river sailing boat, gondola

Duden Sense: 2

CLASS: loud, Venice, music, beat, drums

MEMBER: drums, gondola, cymbal, gong

Duden Sense: 5

Evaluation Results

- Precision:

Source	100 clusters		200 clusters	
No. of clusters	82		91	
Ann 1	59	72%	68	75%
Ann 2	64	78%	65	71%

Inter-annotator agreement: 81/85%

- Recall (from 113 Duden senses):

Source	100 clusters		200 clusters	
Ann 1	46	41%	54	48%
Ann 2	51	45%	52	46%

Summary of Cluster Analyses

- The analyses demonstrated that and how the clusters can be used to predict and discriminate noun sense.
- The predictions are not perfect, but approximately correspond to our linguistic intuitions.
- The predictability of the analyses was evaluated against dictionary definitions and human judgements, precision 71-78%, recall 41-48%.

Conclusions

- Consistent with prior psycholinguistic research, we observed associations to multiple meanings of ambiguous target words.
- Multiple meanings of homonyms were active during both *picture+word* and *word-only* processing, but the degree to which alternative meanings were active was different.
- Semantic associations and association strengths from word-based norming studies do not necessarily generalise for experiments using depicted materials.
- A cluster analysis demonstrated that we can capitalise on the semantic associations and identify and discriminate the various noun senses.