Waar bestaat een onderzoek uit?

- Opstellen van een hypothese of een vraag.
- Bouw een (relevant) experiment.
- Verzamel data.
- “Operationaliseer” data.
- Stel data voor.
- Interpreteer data.

Operationalisering

- Verzamelen van data en op een zodanige wijze voorstellen dat het snel interpreteerbaar is en relevant is voor je hypothese.
- Hierbij is statistiek heel belangrijk.

Voorbeeld

- Vraag: Studie over kansen op arbeidsmarkt van pas afgestudeerden
- Experiment:enquête over het jobs van afgestudeerden.
- Data: representatieve sample van alle studierichtingen, resultaten, leeftijden, …
- Verwerking: statistisch mét foutenmarges.
- Interpretatie: moet steeds genuanceerd zijn

Voorstellen van resultaten

- Kijk eerst hoe je “peers” gelijkaardige experimenten rapporteren.
  - Zo hoef je het warm water niet opnieuw uit te vinden.
  - Zal iedereen je werk begrijpen.
- Beschrijf duidelijk wat je meet en hoe je het beschrijft.
  - Sommige technieken zijn erg standaard en hoeven niet opschreven te worden: standaard afwijking, mutidimensional scaling (MDS), …
  - Anderen vergen uitleg (en soms een voorbeeldje)
    - Hoe berekenen
    - Grenzen
    - Invloed

Experiment in Steels en Belpaeme (2001)

- Artificial simulation of the formation of colour categories and of colour naming under ecological and cultural pressure.

Ingredients of the simulation

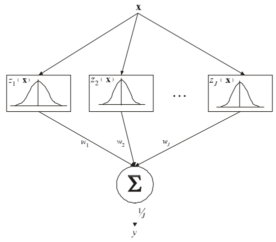
- Population of agents
- Each agent needs:
  - Perception
  - Categorisation
  - Lexicalisation
  - Communication
- Categorisation
  - Learned, through creating and adapting categories
  - Evolved, through mutation of categories

Perception

- Color stimulus in the world reaches the agent as a spectral distribution.
- Needs to be converted into an internal representation faithful to human color perception.
- CIE L*a*b* color space.
  - Mimics human color perception well
  - Inherent opponent nature
  - Is a uniform colour space, with a similarity measure

Categorization

- A color category is represented by an adaptive network



Categorization (2)

- Why adaptive networks?
  - Can represent the prototypical nature of color categories
  - Easy to analyze
  - Allow fast instance-based learning
  - Interpolate between locally tuned units
  - Easily adapted
  - Speedy computation

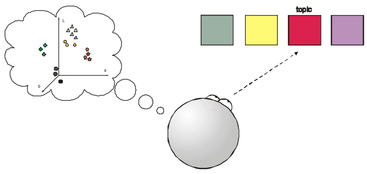
Lexicalization and communication

- A category can be associated with one or more word forms,
  - A form-category association ∈ [0,1]
  - Represents the strength of the association.
- Agents can noiselessly convey word forms to others.

Language games

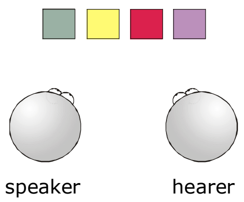
- The dynamics come in the shape of "language games"
  - Steels, Batali, …
- The discrimination game
  - In which an agent creates new categories to discriminate the environment
- The guessing game
  - In which two agents engage in an interaction to convey meaning.

Discrimination game



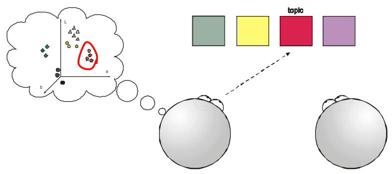
Guessing game

Initialize the game



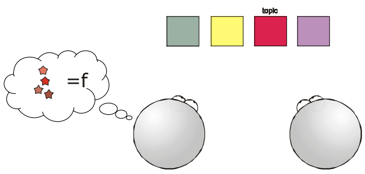
Guessing game

Speaker discriminates topic



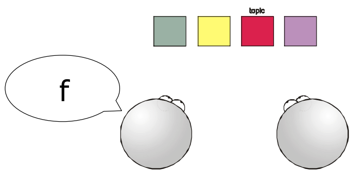
Guessing game

Speaker finds a word form associated with category



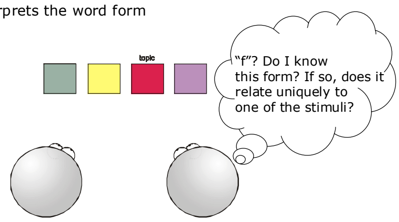
Guessing game

Conveys the word form to the hearer



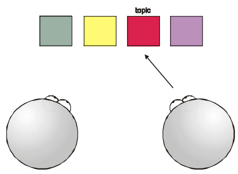
Guessing game

The hearer interprets the word form



Guessing game

The hearer non-verbally points out the topic



Successful game

Failure and adaptation

- Various reasons why guessing game fails
  - Speaker has no discriminating category
  - Speaker has no word form associated with category
  - Hearer lacks the word form in its repertoire
  - Hearer fails to point the topic out.
- Failure and success provide feedback, used by the agent to adapt
  - Create new categories
  - Learn new word forms
  - Adapt categories
    - Add Locally tuned units, change weights of locally tuned units.
  - Strengthen or weaken association between word forms and categories.

Presenting results

- Results have to be presented using clear and objective measures.
- Define measures!
  - Explain
  - Why measure this
  - Formalize
  - Give examples
  - Lower and upper boundaries
  - Error margins

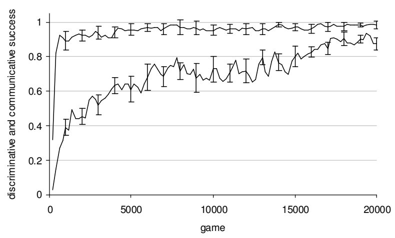
Discriminative success

- How well do agents discrimination? What is the discriminative power of their category sets?
- Needed to see if agents a able to obtain a sufficient and adequate set of categories for discriminating the environment.
- If a discrimination game i is successful: ds_i=1, otherwise ds_i=0
- The discriminative success is the running average of ds of the last T games

$$DS_i = \frac{\sum ds}{T}$$

- Best performace, DS = 1
- Worst performance, DS = 0

Discriminative and communicative success

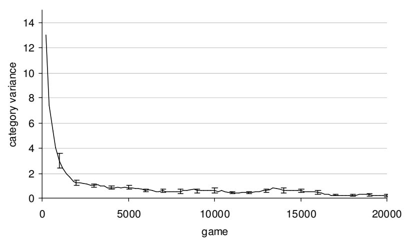


Category variance

- Do categories of different agents agree? Are they coherent?
- Category variance measure

$$cv = \frac{1}{2n(n-1)} \sum_{i=2}^{n} \sum_{j=i-1}^{n} D(A_i, A_j)$$

Category variance



Intra-population category variance

- How well do agents from different populations agree?

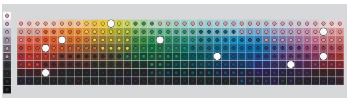
$$cv'(\mathcal{A},\mathcal{A}') = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} D(A_i, A_j)$$

Interesting results

- Nature of the colour categories
- Influence of environmental complexity
- Influence of linguistic communication

Nature of the emerged color categories

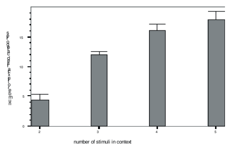
- Example



- No systematicity in the nature of the categories
  - Every run gives a new constellation of categories
- No evolutionary order observed
  - Two categories seems to divide the spectrum in warm/light and cool/dark.
  - Three or more categories do not agree with Berlin and Kay.

Environmental complexity

- Environmental complexity controlled by
  - Number of colour stimuli to be discriminated
  - Similarity of the colour stimuli
- Environment exerts pressure to create more categories.



- Agrees with Rivers' observations: low-technological cultures have fewer colour terms.

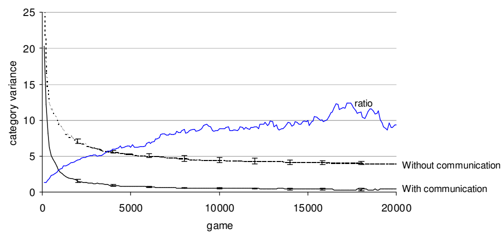
Coherence trough communication

- The linguistic task forces the agents to have identical colour concepts.

  If not, communication would be impossible.

- Agent tune their categories to the categories of others
  - They have no access to each others categories
  - Categories are coupled indirectly through the environment and the lexicon.

Coherence through communication (2)



Conclusions

- Colour categories become **shared** in a population, if the population has the same **ecology** and if the members **communicate**.

- Shared colour categories can be formed through cultural interaction.
  - There is no need for innateness to explain shared color categories in a culture.
  - But… Without strong ecological bias the creation of colour categories is opportunistic and does not follow predictions made by Berlin and Kay.

- In addition… Evolutionary experiments demonstrate weaknesses in universalist theories.