Simplicity and informativeness in the cultural evolution of language

Jon W. Carr, Kenny Smith, Jennifer Culbertson, Simon Kirby

Centre for Language Evolution
School of Philosophy, Psychology and Language Sciences
University of Edinburgh
Pressures shaping language

Language
Pressures shaping language

Simple Language
Pressures shaping language

Simple Language Informative

Learning Communication
Kinship terms are simple and informative

Kinship Categories Across Languages Reflect General Communicative Principles

Charles Kemp and Terry Regier

Languages vary in their systems of kinship categories, but the scope of possible variation appears to be constrained. Previous accounts of kin classification have often emphasized constraints that are specific to the domain of kinship and are not derived from general principles. Here, we propose an account that is founded on two domain-general principles: Good systems of categories are simple, and they enable informative communication. We show computationally that kin classification systems in the world’s languages achieve a near-optimal trade-off between these two competing principles. We also show that our account explains several specific constraints on kin classification proposed previously. Because the principles of simplicity and informativeness are also relevant to other semantic domains, the trade-off between them may provide a domain-general foundation for variation in category systems across languages.

Concepts and categories vary across cultures but may nevertheless be shaped by universal constraints (1–4). Cross-cultural studies have proposed universal constraints that help to explain how colors (5, 6), plants, animals (7–8), and spatial relations (9, 10) are organized into categories. Kinship has traditionally been a prominent domain for studies of this kind, and researchers have described many constraints that help to predict which of the many logically possible kin classification systems are encountered in practice (11–15). Typically these constraints are not derived from general principles, although it is often suggested that they are consistent with cognitive and functional considerations (2, 11–13, 15). Here, we show that major aspects of kin classification follow directly from two general principles: Categories tend to be simple, which minimizes cognitive load, and to be informative, which maximizes communicative efficiency. Principles like these have been discussed in other contexts by previous researchers (16–19). For example, Zipf suggested that word-frequency distributions achieve a trade-off between simplicity and communicative precision (20, 21), Hawkins (22) has suggested that grammars are shaped by a trade-off between simplicity and communicative efficiency, and Rosch has suggested that category systems “provide maximum information with the least cognitive effort” (p. 190 of 23).

Figure 1A shows a simple communicative game that helps to illustrate how kin classification systems are shaped by the principles of simplicity and informativeness. The speaker has a specific relative in mind and utters the category label for that relative. Upon hearing this category label, the hearer must guess which relative the speaker had in mind. Typically the hearer can eliminate some of the possible relatives in the category, and the game ends when the hearer calls out the correct relative. The relative in mind is a sister, so the speaker utters the category “sister.” The hearer can say “mother” to eliminate both brothers and “father” to eliminate both sisters. The game is aimed at maximizing the speaker’s informativeness, and it is presumed that this is the hearer’s goal as well. The hearer will, of course, guess another relative if there is no correct relative in the category (i.e., if the relative in mind is a grandson). To be optimal, the informativeness of the strategy must be maximized, which means that the speaker should choose the category that maximizes her probability of being correct. In this case, the speaker should choose “sister,” which is the label for the relative in mind, because it is the most informative category to use.

Kemp & Regier (2012)
Kinship terms are simple and informative

Kemp & Regier (2012)
Lab experiments

Learning-only
Lab experiments

Learning-only

Communication-only
Lab experiments

Learning-only

Learning + Communication
Kinship terms are simple and informative

Kemp & Regier (2012)
Learning and communication pressures

Informative

Simple
Learning and communication pressures

- Informative
- Simple

Learning-only
Learning and communication pressures

- Informative
- Simple

- Learning-only
- Communication-only
Learning and communication pressures

- Learning-only
- Learning + communication
- Communication-only

Informative vs. Simple
Learning and communication pressures

Kirby, Cornish, & Smith (2008)

Kirby, Tamariz, Cornish, & Smith (2015)
Iterated learning gives rise to informative spatial terms

Language evolution in the lab tends toward informative communication

Alexandra Carstensen (abc@berkeley.edu)
Jing Xu (jing.xu@jhu.edu)
Cameron T. Smith (smithct@berkeley.edu)
Terry Regier (terry.regier@berkeley.edu)

Department of Psychology, Department of Linguistics, Cognitive Science Program, University of California, Berkeley, CA 94720 USA
Department of Neurology, Johns Hopkins University, Baltimore, MD 21287 USA

Abstract

Way do languages-based human experience into categories in the ways that they do? Languages vary widely in their category systems but not arbitrarily, and one possibility is that the cultural histories of languages reflect universal communicative needs. Carstensen with this idea, it has been shown that artificial category systems used to support highly informative communication. However, it is not yet known what process renders these informative systems. Here we show that humans can find cultural transmission in the lab produces systems of semantic categories that converge around greater informativeness, in the domain of color and spatial relations. These findings suggest that large-scale cultural transmission over historical time could have produced the diverse yet informative category systems found in the world's languages.

Keywords: Informativeness, communication, language evolution, human learning, cultural transmission, spatial cognition, color naming, semantic universals.

The origins of semantic diversity

Languages vary widely in their fundamental units of meaning—the concepts and categories they encode in simple words or other basic forms. For example, some languages have a single color term spanning green and blue (Berlin & Kay, 1969), and some have a spatial term that captures the notion of being in water (Levinson & Meira, 2010, 496), and some of which is captured by a single word in English.

Regier’s (2012) kinship study. Levinson (2012) pointed out that although this research explains cross-language semantic variation in communicative terms, it does not tell us “where our categories come from” (p. 969), that is, it does not explain what process gives rise to the diverse systems of informative categorization. Levinson suggested that a possible answer to that question may be in a line of experimental work that explores human simulations of cultural transmission in the laboratory and “shows how categories are based through iterated learning across simulated generations” (p. 969). We agree that prior work explaining cross-language semantic variation in terms of informative communication has not yet addressed this crucial question, and we address it here.

Iterated learning and category systems

The general idea behind iterated learning studies is that of a chain or sequence of learners. The first person in the chain produces some behavior; the next person in the chain observes that behavior, learns from it, and then produces behavior of her own; that learned behavior is then observed by the next person in the chain, who learns from it, and so on. This experimental paradigm is useful as it allows us to investigate how cultural information across generations, the learned behavior, generally changes as it is filtered through the minds of learners.

Iterated learning and related learning studies have been used to explain a wide range of phenomena, including the origins of cultural diversity and the evolution of language.
Iterated learning gives rise to informative spatial terms

Carstensen, Xu, Smith, Regier (2015)
Experiment 1
Stimuli

Angle

Size
## Stimuli

<table>
<thead>
<tr>
<th>Angle</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: The diagram shows a matrix of stimuli varying in angle and size.*
Stimuli

- Angle
- Size
Conditions

- **Angle-only**
  - Easy to learn but low informativeness

- **Size-only**
  - Informative but hard to learn

- **Angle & Size**
  - Informative but hard to learn
Training phase

This is a zix
Test phase

What is this called?  >2$ if correct
reb  zix  wud  pov
Result: Learnability advantage for the less informative systems
 Experiment 2
Iterated learning
Results
Simplicity

Informativeness

![Graph showing changes in complexity and communicative cost over generations.](image)
Measuring simplicity

Categorization Under Complexity: A Unified MDL Account of Human Learning of Regular and Irregular Categories

David Fase
Department of Psychology
Center for Cognitive Science
Rutgers University
Piscataway, NJ 08854
dfase@ruccs.rutgers.edu

Jacob Yehman
Department of Psychology
Center for Cognitive Science
Rutgers University
Piscataway, NJ 08854
jacob@ruccs.rutgers.edu

Abstract

We present an account of human concept learning—i.e., learning of categories from examples—based on the principle of minimum description length (MDL). In support of this theory, we tested a wide range of two-dimensional concept types, including both regular (simple) and highly irregular (complex) structures, and found the MDL theory to give a good account of subjects' performance. This suggests that the intrinsic complexity of a concept (i.e., its description length) systematically influences its learnability.

1 The Structure of Categories

A number of different principles have been advanced to explain the manner in which humans learn to categorize objects. It has been variously suggested that the underlying principle might be the similarity structure of objects [1], the manipulability of decision boundar-ies [2], or Bayesian inference [3][4]. While many of these theories are mathematically well-grounded and have been successful in explaining a range of experimental findings, they have commonly only been tested on a narrow collection of concept types similar to the simple unmodal categories of Figure 1(a)-(b).
Measuring simplicity
Measuring simplicity

4×8

4×8
Measuring simplicity
What makes language informative?

<table>
<thead>
<tr>
<th>Property</th>
<th>Example</th>
<th>Learning</th>
<th>Communication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expressivity</strong></td>
<td>A system of many categories is more informative than a system of few categories</td>
<td><img src="expressivity.png" alt="Example" /></td>
<td>Pressure for simplicity</td>
</tr>
<tr>
<td><strong>Convexity</strong></td>
<td>A system of convex categories is more informative than a system of nonconvex categories</td>
<td><img src="convexity.png" alt="Example" /></td>
<td></td>
</tr>
</tbody>
</table>
What makes language informative?

**Expressivity** A system of many categories is more informative than a system of few categories.

**Convexity** A system of convex categories is more informative than a system of nonconvex categories.
What makes language informative?

**Expressivity** A system of many categories is more informative than a system of few categories.

**Convexity** A system of convex categories is more informative than a system of nonconvex categories.
What makes language informative?

**Expressivity** A system of many categories is more informative than a system of few categories.

**Convexity** A system of convex categories is more informative than a system of nonconvex categories.
What makes language informative?

**Expressivity** A system of many categories is more informative than a system of few categories.

**Convexity** A system of convex categories is more informative than a system of nonconvex categories.
What makes language informative?

**Expressivity** A system of many categories is more informative than a system of few categories.

**Convexity** A system of convex categories is more informative than a system of nonconvex categories.
What makes language informative?

**Property**

**Expressivity** A system of many categories is more informative than a system of few categories.

**Convexity** A system of convex categories is more informative than a system of nonconvex categories.
What makes language informative?

**Property**

**Expressivity** A system of many categories is more informative than a system of few categories.

**Convexity** A system of convex categories is more informative than a system of nonconvex categories.

---

**Example**

[Diagram showing examples of expressivity and convexity]

---

**Learning**

Pressure for simplicity

---

**Communication**

Pressure for informativeness

---

Carstensen et al.
Conclusions

The pressure from learning yields simple languages, but this can manifest itself through multiple semantic properties:

**Expressivity** Loss of words/categories to aid learning

**Convexity** Reorganization of the meaning space to aid learning

The pressure from communication has the same effect on convexity as learning does, making it difficult to identify the causal mechanism.

In iterated learning, the pressure for simplicity favours semantic category systems that are convex – informativeness comes along for the ride.
Vielen Dank!