



Sentence and Word Complexity

Jeffrey Heinz and William Idsardi Science **333**, 295 (2011); DOI: 10.1126/science.1210358

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by clicking here.

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines here.

The following resources related to this article are available online at www.sciencemag.org (this information is current as of November 12, 2015):

A correction has been published for this article at: http://www.sciencemag.org/content/333/6051/1825.1.full.html

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

http://www.sciencemag.org/content/333/6040/295.full.html

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

http://www.sciencemag.org/content/333/6040/295.full.html#related

This article **cites 15 articles**, 3 of which can be accessed free: http://www.sciencemag.org/content/333/6040/295.full.html#ref-list-1

This article has been **cited by** 2 articles hosted by HighWire Press; see: http://www.sciencemag.org/content/333/6040/295.full.html#related-urls

This article appears in the following **subject collections**: Neuroscience

http://www.sciencemag.org/cgi/collection/neuroscience

the fruit fly intestine, the hindgut, starts out as a simple tube along the embryo's midline. Through a process that occurs without cell division, this tube first dips ventrally (toward the fly's "belly"), and then rotates leftward by 90° to create a net rightward bend. Seeking a cellular basis for the rotation, Taniguchi et al. make the key observation of a small statistical bias in hindgut cell shapes with respect to the embryo's left-right axis. Cell-cell boundaries that make angles between -90° and 0° with the tube's long axis (left boundaries) appear more frequently than boundaries that make angles between 0° and 90° (right boundaries). The authors call this pattern planar cell-shape chirality (PCC). They identify the cell-cell adhesion molecule Drosophila E-cadherin (DE-cadherin) as a factor required for both PCC and gut rotation, showing that it is preferentially enriched on left boundaries. Mutations in a motor protein involved in intracellular movement, known as unconventional myosin ID (MyoID), reverse the polarity of DE-cadherin accumulation and PCC. This is consistent with MyoID's previously identified role in setting the direction of gut rotation (5, 6).

Because mutations in DE-cadherin cause all cell boundaries to expand, Taniguchi et al. suggest that DE-cadherin limits boundary expansion by increasing boundary tension. They propose that left-biased tension is sufficient to produce a leftward tissue rotation. Indeed, computer simulations identify one possible mechanism by which this might work. First, left-biased tension drives cell shape change and rearrangement while the endpoints of the tube remain fixed. Then, rotation occurs in the absence of asymmetric tension, and the tube twists as the cells relax back toward more regular shapes.

Together, Tang et al. and Taniguchi et al. highlight how statistical differences in cell behavior across a large population can lead to stereotyped, tissue-level morphogenesis. They also highlight several key ways in which mathematical models provide an essential predictive bridge between cell- and tissuelevel dynamics. In the mouse lung, it is intuitively clear that biases in cell division orientations could cause differential increases in tube length versus circumference, and previous work had shown that oriented cell divisions can contribute to tube shape (7-10). A model, however, was essential to show quantitative sufficiency. In the case of the fruit fly hindgut, it is far from obvious how biasing tension on left boundaries will produce a leftward twist. Here, mathematical models step in when intuition fails, and provide plausible testable hypotheses.

For both systems, the mathematical models provide a framework for exploring the molecular mechanisms that control local cell polarity and coordinate its tissue-wide effects. One obvious candidate in both cases is the signaling pathway known as the planar cell polarity (PCP) pathway, which controls cell division orientations and cellular polarities in many other contexts (11, 12). In the experiments conducted by Tang et al. and Taniguchi et al., however, disrupting PCP function had no effect on these developmental processes, suggesting that other mechanisms are at work.

In the mouse lung, a key question is: How does ERK signaling shape the distribution of cell division angles? The nature of the wildtype distribution suggests that cells partition between two qualitatively distinct orientation states: strictly longitudinal or random. Tang et al. hypothesize that the longitudinal state is the default, that ERK signaling overrides this default to randomize division axes, and that Spry1/2 tune ERK signaling to achieve a balance between longitudinal and random divisions. But how does a graded change in ERK levels control the fraction of cells that inhabit these two states? Does ERK signaling merely gate the response to a longitudinal cue, or does it directly control a transition between distinct phenotypic states?

Likewise, the Taniguchi et al. study provides a starting point for thinking about how local left-right asymmetries in force generation could drive chiral rotation, but how do

these asymmetries arise? The observation that MyoID mutants exhibit reversed PCC and gut rotation implies an intrinsic mechanism for breaking chiral symmetry that can be biased in either direction. The genetic requirements for DE-cadherin and MyoID suggest that symmetry breaking occurs shortly before hindgut rotation and requires local interaction across cell-cell boundaries. By contrast, in vertebrates, establishment of left-right asymmetry occurs far in space and time from the organs undergoing chiral morphogenesis, which suggests that it may be easier to identify the mechanisms involved.

These studies signal a growing trend in which classical molecular and genetic approaches merge with quantitative microscopy, image analysis, and modeling to provide new insights into the cellular dynamics of tissue morphogenesis. It is likely, however, that we are seeing just the tip of an iceberg.

- 1. B. Lubarsky, M. A. Krasnow, Cell 112, 19 (2003).
- 2. N. Tang et al., Science 333, 342 (2011).
- 3. K. Taniguchi et al, Science 333, 339 (2011).
- 4.]. M. Mason et al., Trends Cell Biol. 16, 45 (2006).
- 5. L. A. Baena-Lopez et al., Curr. Biol. 15, 1640 (2005).
- 6. M. Matsuyama et al., PLoS Genet. 5, e1000427 (2009).
- 7. S. M. Meilhac et al., J. Cell Biol. 164, 97 (2004). 8. E. Fischer et al., Nat. Genet. 38, 21 (2006).
- 9. S. Hozumi et al., Nature 440, 798 (2006).
- 10. P. Spéder, G. Adám, S. Noselli, Nature 440, 803 (2006).
- 11. H. Song et al., Nature 466, 378 (2010).
- 12. M. Segalen, Y. Bellaïche, Semin. Cell Dev. Biol. 20, 972

10.1126/science.1209687

PSYCHOLOGY

Sentence and Word Complexity

Jeffrev Heinz¹ and William Idsardi²

Do humans learn the sentence and sound patterns of natural languages through distinct learning mechanisms?

ur understanding of human learning is increasingly informed by findings from multiple fields—psychology, neuroscience, computer science, linguistics, and education. A convergence of insights is forging a "new science of learning" within cognitive science, which promises to play a key role in developing intelligent machines (1, 2). A long-standing fundamental issue in theories of human learning is whether there are specialized learning mechanisms for cer-

¹Department of Linguistics and Cognitive Science, University of Delaware, 42 East Delaware Avenue, Newark, DE 19716, USA. ²Department of Linguistics and Program in Neuroscience and Cognitive Science, University of Maryland, 1401 Marie Mount Hall, College Park, MD 20742, USA. E-mail: heinz@udel.edu

tain tasks or spheres of activity (domains). For example, is learning how to open a door (turning the handle before pulling) the same kind of "learning" as putting up and taking down scaffolding (where disassembly must be done in the reverse order of assembly)? Surprisingly, this issue plays out within the domain of human language.

Language perception is organized at different levels, each with its own internal organizing principles: the organization of sounds into words (phonology), the organization of roots and affixes into words (morphology), and the organization of words into phrases into sentences (syntax). Are there any differences among the patterns observed at each level? And if there are, are specialized or

Theoretical computer science provides a mathematically rigorous way to characterize patterns in terms of strings, i.e., sequences of more basic units. For example, sentences are sequences of words; words are sequences of sounds. Patterns in any domain can be described as a set of strings or a probability distribution over strings. The Subregular and Chomsky hierarchies (3) arrange all logically possible patterns into nested regions of complexity (see the figure). These regions have multiple, convergent definitions which distill the necessary properties of devices (such as grammar, in the case of language) that can recognize or generate the strings comprising the pattern (or, in the case of distributions, their probabilities). This mathematical framework permits the comparison of patterns and their complexity across different levels or domains, linguistic or otherwise.

Every natural language distinguishes well-formed and ill-formed sentences and words. Every sentence in this article is wellformed; ill-formed sentences are obtained by shifting every period one word to the left. Likewise, English words must contain proper sequences of English sounds (4). For example, "gdansk" and "srem" are not possible native English words, though they are possible words in other languages. English speakers can readily assent to new coinages ("bling") while avoiding others ("gding").

An important, but perhaps overlooked, difference is that sound patterns are less complex than sentence patterns according to the Subregular and Chomsky hierarchies. Capturing which patterns of words are wellformed sentences can require context-free or even mildly context-sensitive computations (5). Thus, similar to the assembly and disassembly of scaffolding, English exhibits recursion, because sentences ("The mouse ran away") can be contained in larger ones ("The mouse that the cat chased ran away"). The recursive nature of sentence patterns is a defining characteristic of natural language (6), and demonstrably makes such sentence patterns at least context-free.

By contrast, sound patterns are measurably simpler because identifying wellformed patterns of sounds does not require context-free computations (7). One kind of sound pattern restricts adjacent sounds (such

Multiple, distinct Single language language learners learner D. B. **Finite** Strictly Local Regular Contextsensitive Mildly contextsensitive Contextsensitive Recursively enumerable

Complexity hierarchies for patterns. The Subregular and Chomsky hierarchies divide all logically possible patterns into nested regions (4). A is a pattern of permissible consonant clusters in English (14); B is the long-distance sound pattern of Samala (9); C is the recursive pattern of sentences within sentences in English; D is a particular sentence pattern in Swiss German (5).

as "gding"). Another kind of sound pattern restricts sounds over long distances (8). For example, Samala, a language native to California, does not allow words containing both "s" and "sh." Consequently, there are words such as "shtoyonowonowash" (which means "it stood upright") but none like "shtoyonowonowas" (9). Moreover, all sound patterns fall into the "regular" region (less restricted) of the Chomsky hierarchy (7), and probably belong to even less complex regions (10).

What are the possible explanations for this computationally measurable difference between sentence patterns and sound patterns? One possibility is that sound patterns are constrained by the human nervous, motor, and auditory systems in ways that sentence patterns are not. That is, the momentto-moment configurations of the vocal tract constrain the moment-to-moment articulation of sounds. However, long-distance sound patterns cannot be due to articulation alone because the tongue does not retain the "sh" posture throughout "shtoyonowonowash."

Because humans do learn languages, another possibility is that the properties of sentence patterns and sound patterns reflect properties of how they do this. It follows that if humans employ distinct learning mecha-

nisms for phonology and syntax, the complexity differential could be explained. There is a convergence of results from philosophy, psychology, and computer science (1, 11-13), showing that learning is only possible if learners (humans or machines) are restricted in the generalizations they are allowed to consider. Indeed, the successes in the "new science of learning" carefully tailor hypothesis spaces so that learners succeed with reasonable amounts of data and effort. In fact, recent, distinct computational models for learning sound patterns (10, 14) and sentence patterns (15, 16) succeed because they exclude certain patterns from consideration. For instance, algorithms for soundpattern learning exclude context-free patterns. In this way, these results demonstrate the utility of multiple, differentiated algorithms for language learning.

Task-specific knowledge and its acquisition are familiar in robotics (17) and biology (18). Human and robot locomotion, for example, are often modeled with systems with less than context-free power (19). Thus, modular learning proposals may be based not on the content of the domains (language versus locomotion), but rather on the informational complexity of the patterns in each domain.

Although single, general-purpose language-learning models (20) cannot be ruled out, they face a series of challenges. Such models predict that no complexity differential between sound patterns and sentence patterns should exist: Any sentence pattern ought to be a possible sound pattern and vice versa. Moreover, such unitary algorithms must enable the learning of both sound patterns and sentence patterns from reasonable amounts of data. Finally, these models should also offer some explanation for the observed difference in complexity between sentence and sound patterns (or disprove it).

While complexity differentials do not entail distinct learning mechanisms, the hypothesis that humans learn sound patterns separately from sentence patterns provides a viable explanation for the difference in complexity observed between them. The utility of specialized or general learning mechanisms (or perhaps even both) for language learning will become clearer through further empirical collaborations. Indeed, psychologists and linguists are currently testing these hypotheses with artificial language-learning experiments (21), an approach that can determine $\frac{\pi}{2}$

whether people make the same kind of generalizations when they encounter both words and sentences.

References and Notes

- 1. J. B. Tenenbaum et al., Science 331, 1279 (2011).
- 2. A. N. Meltzoff et al., Science 325, 284 (2009).
- 3. R. C. Berwick et al., Trends Cogn. Sci. 15, 113 (2011).
- 4. M. Halle, Linguistic Theory and Psychological Reality (MIT Press, Cambridge, MA, 1978).
- 5. S. Shieber, *Linguist. Philos.* **8**, 333 (1985).
- 6. M. D. Hauser, N. Chomsky, W. T. Fitch, Science 298, 1569

- 7. R. Kaplan, M. Kay, Comput. Linguist. 20, 331 (1994).
- 8. S. Rose, R. Walker, *Language* **80**, 475 (2004).
- 9. R. Applegate, Samala-English Dictionary: A Guide to the Samala Language of the Ineseño Chumash People (Santa Ynez Band of Chumash Indians, Santa Ynez, CA, 2007).
- 10. J. Heinz, Linguist. Inq. 41, 623 (2010).
- 11. V. Vapnik, Statistical Learning Theory (Wiley, New York, 1998).
- 12. M. A. Nowak et al., Nature 417, 611 (2002).
- 13. C. de la Higuera, Grammatical Inference: Learning Automata and Grammars (Cambridge Univ. Press, Cambridge, 2010).
- 14. B. Hayes, C. Wilson, Linguist. Inq. 39, 379 (2008).
- 15. L. Becerra-Bonache, J. Case, S. Jain, F. Stephan, Theor.

- Comput. Sci. 411, 2741 (2010).
- 16. A. Clark, S. Lappin, Linguistic Nativism and the Poverty of the Stimulus (Wiley-Blackwell, Chichester, UK, 2011).
- 17. C. Belta et al., IEEE Robot. Autom. Mag. 14, 61 (2007).
- 18. C. R. Gallistel, The Organization of Learning (MIT Press, Cambridge, MA, 1990).
- 19. W. I. Sellers et al., J. Anat. 204, 343 (2004).
- 20. N. Chater, P. Vitányi, J. Math. Psychol. 51, 135 (2007).
- 21. V. Folia et al., Lang. Learn. 60 (suppl. 2), 188 (2010).
- 22. This work was supported by the NSF (grant 1035577 to J.H.) and the National Institute on Deafness and Other Communication Disorders (grant 7R01DC005660-07 to W.I.).

10.1126/science.1210358

CHEMISTRY

Pores Within Pores—How to Craft Ordered Hierarchical Zeolites

Karin Möller and Thomas Bein

Veolites are aluminosilicate crystals that have internal networks of angstrom-size pores, similar to the dimensions of small molecules. They are among the most widely used materials in heterogeneous catalysis (1, 2) because of their defined structure and composition. Although zeolites are very potent solidacid catalysts, their catalytic applications have been limited to processing smaller molecules; their internal pores are not readily accessed by molecules exceeding 1 nm in size. Major efforts have been directed to overcoming this limitation. On page 328 of this issue, Na et al. (3) present a new strategy for creating thin zeolite walls, containing small pores, that grow into structures forming larger pores that can catalyze reactions with larger molecules.

To date, two major synthesis strategies have been explored to create zeolites with additional larger pores. One is to form a secondary pore system of larger size than the zeolitic micropores within the zeolite crystal, thereby allowing faster diffusion of larger molecules into the zeolite particles. In many zeolite synthesis routes, molecular "templates" are added to aid the growth of the aluminosilicate crystals; when the synthesis is completed and the template molecules are removed, zeolitic pore spaces remain. Traditional methods in zeolite synthesis use a single molecular template with a size similar to or smaller than the micropore dimensions. This micropore size bar-

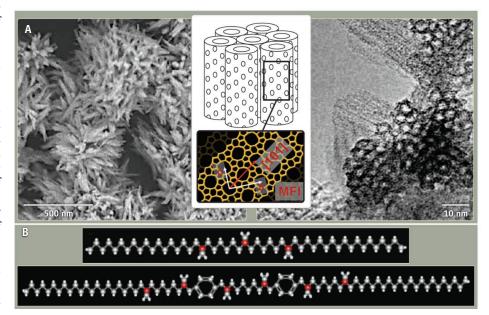
Department of Chemistry, Ludwig-Maximilians University, 81377 Munich, Germany. E-mail: karin.moeller@cup.unimuenchen.de; bein@lmu.de

rier was broken in the early 1990s with the synthesis of mesoporous oxides by using larger amphiphilic surfactant templates (4). However, the local order in the resulting (alumino-)silicates was lost, and with it their strong acidity. Efforts to recrystallize these amorphous walls into zeolitic structures were usually unsuccessful.

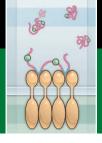
Dual-templating approaches are now being explored that combine the advantages of zeolites and the mesoporous oxides. In addition to the molecular zeolite templates, larger assemThin walls of crystalline zeolites can be assembled into hexagonal nanopore networks, which expands the range of their catalytic reactions to larger molecules.

blies such as the above surfactant micelles (5), polymers (6), or objects such as carbon beads or fibers (7) were used to create mesoporosity. Alternative approaches include chemical treatments that partially dissolve the crystalline zeolite lattice and create larger intraparticle cavities (8). The need for a secondary template can also be avoided by the direct assembly of nanosized zeolite particles, thus creating mesoporous interparticle voids (9).

The other approach for allowing access of larger molecules to zeolite pores is to pre-



Ordering the walls. Na et al. have used large molecular templates to grow thin walls of zeolites into morphologies that create ordered mesopores. (A) Two electron microscopy images of the hexagonally grown mesoporous MFI-type zeolite with extremely high surface area. The inset shows a schematic of the hierarchical structure and the MFI pore framework of the zeolite walls. (B) Examples of the bifunctional templating molecules used in the synthesis that bear ammonium groups and long alkyl chains (white spheres, hydrogen; gray spheres, carbon; red spheres, nitrogen).



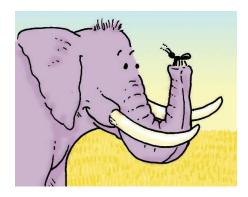
Stress sensor engagement

1830



SPORE Prize Essay

1838



The Elephant is beautiful, muscular, and respected by all the animals in the jungle. The Ant is small and ignored. Could an Ant possibly build a strong relationship with an Elephant? This Ant was going to try.

First, the Ant sought national authorization to conduct a clinical research project that would develop baseline studies and collect baseline data necessary for future clinical trials. After submitting the research protocol, the Ant waited 15 long months for approval by Congo's only Institutional Ethics Committee, and two more months for authorization from the Ministry of Health. A 17-month delay could compromise the rest of the project, thought the Ant with alarm. The work plan was often misunderstood, and as the Ant explained again and again how the money would be used to address specific challenges, she worried that the Elephant would move on and leave her behind.

The Ant realized that a good research team must be multidisciplinary, consisting of junior and seniors scientists selected by an experienced panel from a list of qualified applicants. This would be a challenge in a place with limited postgraduate academic opportunities. To overcome this limitation,

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the past 3 months or matters of general interest. Letters are not acknowledged upon receipt. Whether published in full or in part, Letters are subject to editing for clarity and space. Letters submitted, published, or posted elsewhere, in print or online, will be disqualified. To submit a Letter, go to www.submit2science.org.

the Ant launched an open call for applications. The other animals in the jungle viewed the Ant's new approach with suspicion.

To invest in infrastructure, the Ant renovated an abandoned facility into the first molecular biology laboratory of the Faculty of Health Sciences, and then equipped it. Now the other animals started to appreciate the Ant's hard work. They congratulated her for the change and encouraged her to maintain the spirit.

To create a culture of research, the Ant had to be thoughtful and innovative. She stimulated scientific discussions by implementing regular scientific meetings. But how would she attract students and scientists to these meetings and foster interest and loyalty? The Ant formed brigades of students

to urge others to participate. A year later, the seminar room was always filled with an enthusiastic audience.

Once she had met these challenges, the Ant invited the Elephant to her home to share a cup of tea. She told him about all of her accomplishments, and showed him the new facilities. When the Elephant returned home, he was smiling and convinced. And he wondered, "What kind of Ant is this, this Ant who acts like an Elephant?"

It was just as the Ant had hoped. Next, the Ant hopes to sustain positive momentum and establish stable local research teams that will regularly publish in international scientific journals (2). The moral of the story: For young Congolese scientists wondering how to contribute scientifically to their country, the metamorphosis from tiny Ant to majestic Elephant is possible, but it will require time, cunning, and determination.

Congolese Foundation for Medical Research/Faculty of Health Sciences, University Marien Ngouabi, Brazzaville, Congo, and Institute for Tropical Medicine, University of Tübingen, Germany. E-mail: fntoumi@fcrm-congo.com

References

- 1. E. Dolgin, Nat. Med. 16, 8 (2010).
- 2. F. Ntoumi, G. Priebe, Malaria J. 9 (suppl. 3), S7 (2010).

CORRECTIONS AND CLARIFICATIONS

Editors' Choice: "Loud enough?" by J. S. Yeston (12 August, p. 803). Dalian's location should have been specified as Northeast, rather than Northwest, China.

News Focus: "Climate change sparks battles in classroom" by S. Reardon (5 August, p. 688). The credit for the second image (bottom of page 688) was incorrect. The correct credit is Morgan Heim, Cooperative Institute for Research in Environmental Sciences. The credit has been corrected in the HTML version online.

Reports: "The plant cell wall—decomposing machinery underlies the functional diversity of forest fungi" by D. C. Eastwood *et al.* (5 August, p. 762). The second sentence of the caption for Fig. 2C should read, "Black trace, nitrogen-rich medium (+N); red trace, nitrogen-depleted minimal medium (-N)."

Perspectives: "Sentence and word complexity" by J. Heinz and W. Idsardi (15 July, p. 295). Due to a production error, the inner region in the figure was mislabeled "Context-sensitive." It should read "Context-free."

TECHNICAL COMMENT ABSTRACTS

Comment on "The Response of Vegetation on the Andean Flank in Western Amazonia to Pleistocene Climate Change"

Surangi W. Punyasena, James W. Dalling, Carlos Jaramillo, Benjamin L. Turner

Cárdenas *et al.* (Reports, 25 February 2011, p. 1055) used the presence of *Podocarpus* pollen and wood to infer ≥5°C cooling of Andean forests during Quaternary glacial periods. We show that (i) *Podocarpus* has a wide elevation range in the Neotropics and (ii) edaphic factors cannot be discounted as a factor governing its distribution. Paleoecologists should therefore reevaluate *Podocarpus* as a cool-temperature proxy.

Full text at www.sciencemag.org/cgi/content/full/333/6051/1825-b

Response to Comment on "The Response of Vegetation on the Andean Flank in Western Amazonia to Pleistocene Climate Change"

Macarena L. Cárdenas, William D. Gosling, Sarah C. Sherlock, Imogen Poole, R. Toby Pennington, Patricia Mothes

Punyasena *et al.* question our interpretation of climate-driven vegetation change on the Andean flank in western Amazonia during the middle Pleistocene and suggest that the use of *Podocarpus* spp. as a proxy of past climate change should be reassessed. We defend our assertion that vegetation change at the Erazo study site was predominantly driven by climate change due to concomitant changes recorded by multiple taxa in the fossil record. Full text at www.sciencemag.org/cgi/content/full/333/6051/1825-c

