in the pieces of the landscape shown in pink. Their method is directly related to the time-resolved probing of vibrations by means of resonance Raman spectroscopy, as previously reported by Mathies and co-workers (*7*, *8*). The main difference is that Takeuchi *et al*. probe the vibrational motions in time rather than frequency, which has the practical result of lowering the frequency of the vibrations that can be accessed. Thus, they can observe the evolution of a vibrational motion of the carbon-carbon bond framework at frequencies around 200 cm^{-1} during the isomerization reaction, which represents nuclear motions with a period of 1.6×10^{-13} s.

This lower-frequency motion is similar in time scale to the motion along the reaction coordinate, and the curvature of the potential energy surface along these two dimensions is comparable. It provides details of how the phenyl rings move and twist as they settle into the extended trans conformation, which previously was viewed as a spectator to the motion rather than as part of the action. The complementary computational study in the report highlights the necessity of combining theory and experiment when mapping out these potential energy surfaces.

The report by Takeuchi *et al*. adds to our understanding of a specific class of chemical reactions by providing a new perspective on a model photoisomerization. Their study takes us beyond the question of "how fast" and to the more demanding question of "which way" at the level of the entire molecule. Although the method presented is technically demanding, it could be applied to a wide variety of photoinitiated reactions, including those that take place in complex environments such as proteins.

References
1. S. Takeuchi *et al.*, *Science* **322**, 1073 (2008).
2. C. Dugave, L. Demange, *Chem. Rev.* **103**, 2475 (2003).
3. M. V. der Horst, K. Hellingwerf, *Acc. Chem. Res.* **37**, 13 (2004).
4. W. Heitler, F. London, *Z. Phys.* **44**, 455 (1927).
5. H. Eyring, M. Polanyi, *Z. Phys. Chem. Abt. B* **12**, 279 (1931).
6. G. R. Fleming, *Chemical Applications of Ultrafast Spectroscopy* (Oxford Univ. Press, Oxford, 1986).
7. P. Kukura *et al.*, *Science* **310**, 1006 (2005).
8. P. Kukura *et al.*, *Annu. Rev. Phys. Chem.* **58**, 461 (2007).

BEHAVIOR

A Biolinguistic Agenda

Neurobiology and genetics are helping to generate insights about the evolution of language.

Marc D. Hauser[1] and Thomas Bever[2]

When we transform thoughts into speech, we do something that no other animal ever achieves. Children acquire this ability effortlessly and without being taught, as though discovering how to walk. Damage to specific areas of the brain that are critical to language shows the profound selectivity of cerebral organization, underlining the exquisite biological structure of language and its computational features. Recent advances bring new insights into the neurogenetic basis of language, its development, and evolution, but also reveal deep holes in our understanding.

There are about 7000 living languages spoken in the world today, characterized by both exceptional diversity as well as significant similarities. Despite many controversies in the field, many linguistic scholars generally agree on two points (*1*–*8*). Language as a system of knowledge is based on genetic mechanisms that create the similarities observed across different languages, culturally specific experience that shapes the particular language acquired, and developmental processes that enable the growth and expression of linguistic knowledge. Also, the neural systems that allow us to acquire and process our knowledge of language are separate from those underlying our ability to communicate.

To fulfill a biolinguistic agenda—study of the computational systems inherent to language—we must address the rules and constraints that underlie a mature speaker's knowledge of language; how these rules and constraints are acquired; and whether they are mediated by language-specific mechanisms. We also need to distinguish which rules and constraints are shared with other animals and how they evolved, and to ask how knowledge of language is used in communicative expressions.

There has been little research linking the formal linguistic principles that describe the mature speaker's knowledge of language to the evolutionary, neurobiological, and developmental factors that lead to their instantiation in the adult mind. These principles include computational devices such as hierarchies and dependencies among syntactic categories (e.g., the relationship between determiners such as "the" and "a" followed by nouns), recursive and combinatorial operations, and movement of parts of speech and phrases (e.g., to create a question, many languages move constructions such as "what" or "where" to the front of the sentence). This gap is slowly narrowing, but the separation remains great. It is thus important to clarify the appropriate targets of analysis. In particular, examination of the evolutionary, neurobiological, and developmental aspects of language often focuses narrowly on speech, or in some cases, on the separate issue of communication. Instead, these aspects should be considered in light of the principles discussed, helping to align formal approaches to linguistics with the biological sciences.

Formal approaches to examine linguistic structure are marked by disagreement about the necessary or sufficient computations required to create the expressed languages of the world. Some linguists argue that linguistic form relies on abstract, generative operations that allow phrases and sentences (syntactic structures) to interface with meanings (the semantic system) to create a categorization (lexical terms) in which single words and groups of words convey a specific meaning. Such lexical terms then interface with speech sounds (phonology) to create expressed words in speech or sign. Language has been suggested as an optimal solution to the syntactic-semantics interface, achieved by a small number of computational operations. By comparison, current evolutionary models suggest that the variation in animal body form can be explained by different activation patterns for a few master genes during development. The corresponding idea in linguistics is that the cross-cultural variation in expressed human languages can be explained by a universal set of mental operations, some specific to language, others shared across domains including music, mathematics, and morality (*4*, *9*).

Comparative evolutionary studies suggest that birds, rodents, and primates compute some components of human grammatical

[1]Department of Psychology, Human Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [2]Departments of Linguistics, Psychology and Cognitive Science Program University of Arizona, Tuscon, AZ 85721, USA. E-mail: mdh@wjh.harvard.edu

competence, but cannot attach this capacity to their own communication systems (*10–12*). For example, birds and primates can compute a first-degree finite state grammar, where elements in a string of sounds have specific orders, each predicted by simple statistical associations. This grammar is one of the simplest within a hierarchy of computational operations of increasing complexity and expressive power (*10, 13*). The biggest puzzle, however, is why nonhuman animals cannot integrate these computational capacities with their capacity to communicate. So, although songbirds can combine different notes into a variety of songs, they don't integrate this combinatorial capacity with conceptual abilities to create sounds with varied meaning. Understanding what neural connections are absent, or poorly developed, may help account for this evolutionary bottleneck, and explain why human infants readily produce an infinite variety of meaningful expressions.

Damage to Broca's area and Wernicke's area in the human brain results in distinct patterns of language loss, suggesting that properties of the neocortex make language unique to humans. Artificial language studies show that these cortical areas execute the computations that obey language universals (the principles accessed by all languages, such as specific word orders), but other brain areas are also activated by these computations (*14, 15*). In fact, different cortical areas may compute different kinds of grammars, but such localization does not provide insight into linguistic theories aimed at uncovering principles that guide the mature state of language competence and its acquisition during development.

Does language have its own dedicated brain circuits, or is much or all of this circuitry shared across domains (such as music and language)? For example, language and music rely on hierarchical representations, make use of combinatorial and recursive computations, and generate serially represented structures. But does each domain recruit a general-use ensemble of these processes or does each domain have its own set of processes? Further studies of selective brain damage and brain-imaging experiments should be informative.

Genes associated with particular linguistic deficits can help pinpoint the molecular basis for language, and link issues in evolution with those in development. Yet, we are far from understanding how normal genes are associated with linguistic features. When the gene *FOXP2* was linked to families with a particular language deficit, it seemed that genomics might account for linguistic structure. But the

relationship between *FOXP2* and language turns out to be weak. For example, *FOXP2* exists in songbirds and echo-locating bats; although songbirds have richly structured sound systems that might be properly characterized by a finite state grammar, such grammars are not hierarchically structured, lack syntactic categories (e.g., nouns and determiners), and do not productively generate meaningful variation. Further, the disorders associated with *FOXP2* in humans include articulatory disabilities and are not clearly syntactic, semantic, or computational (*16, 17*). The weak connection between *FOXP2* and these aspects
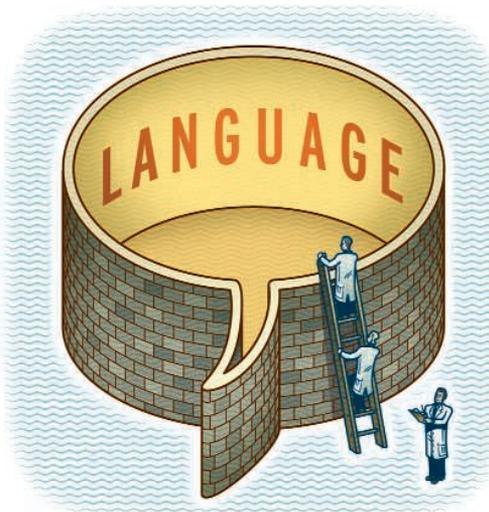
of language should not, however, come as a surprise given that most gene-phenotype relationships involving complex phenotypes (such as language) are weak. Nonetheless, by breaking language down into its component parts and finding potential homologs in other animals (especially those that can be genetically manipulated), we may better understand the evolution, development, and neurobiological breakdown of linguistic function.

Current research on hemispheric lateralization (division of the brain into left and right halves) and language acquisition provides one example of how interdisciplinary work relates to specific theories in linguistics. All right-handed people have strong left-hemisphere lateralization of syntactic function. However, classic investigations of aphasia—the inability to produce or comprehend language—reveal that familially "mixed" right-handers (right-handers with left-handed family members) show more right-hemisphere involvement in language than pure right-handers (*18, 19*). Thus, in familially mixed right-handers, the right hemisphere's involvement in language may be specific to lexical representations (*20*).

Familially mixed right-handers access individual words more readily than global sentence structure, whereas the reverse is true of familially pure right-handers (*21*). Their critical period for language learning is also earlier than that of familially pure right-handers (*22*), which suggests that mixed right-handers are more likely to base their language learning on the acquisition of words as opposed to syntactic structure. These findings are supported by brain-imaging research showing that familially pure right-handers have left-hemisphere activation during lexical access, whereas familially mixed right-handers show more bilateral hemisphere activation (*23*). At the same time, all subjects show left-hemisphere activation for syntactic processes. This confirms the basic hypothesis that mixed right-handers have more distributed representations of lexical knowledge.

What are the implications of such population-level differences in lexical use, access, and representation for linguistic theory? In recent decades, syntacticians have struggled with the role of the lexicon in syntactic architectures. Proposals range from the traditional view that the lexicon is distinct from the computations of syntax, to the view that syntax itself is driven by lexical structures. The observed variability in how the lexicon is accessed and represented suggests that it is indeed a biologically separable component of linguistic knowledge.

Brain imaging, genomics, and new methods for comparative studies have provided the means for better understanding the shared and uniquely human components of language. As some linguists argue, the variation in linguistic form among the world's languages may be as superficial as the variation in animal body forms. The superficiality arises, in each case, because of universal computations that provide the necessary suite of developmental programs to generate the variation. As the biolinguistic agenda advances, however, new generations of linguists will be required to translate their formalisms into testable experiments by biologists and psychologists. For example, language deploys recursive operations and generates hierarchical representations with specific configurations. It is not yet clear how to design experiments to test whether nonlinguistic organisms can acquire these representations, or what factors limit either their acquisition or implementation into communicative expression. Conversely, psychologists and biologists will need to be sensitive to the limitations of their methods and the extent to which they can test linguistic theories. Thus, neuropsychological studies showing deficits in language need

to be accompanied by comparable tests in non-linguistic domains to show that they are language-specific deficits. And studies using brain imaging must acknowledge that localization of function does not provide explanatory power for the linguist attempting to uncover principles underlying the speaker's knowledge of language. These cautions aside, the biolinguistic approach is clearly benefiting from modern technologies to advance our knowledge of what language is, how it is represented, and where it came from.

References

1. M. A. Arbib, *Behav. Brain Sci.* **28**, 105 (2005).
2. E. Bates, *Discuss. Neurosci.* **10**, 136 (1994).
3. T. G. Bever, in *Cognition and Language Development*, R. Hayes, Ed. (Wiley, New York, 1970), pp. 277–360.
4. N. Chomsky, *Linguist. Inq.* **36**, 1 (2005).
5. T. W. Deacon, *The Symbolic Species: The Coevolution of Language and the Brain* (Norton, New York, 1997).
6. R. Jackendoff, *Foundations of Language* (Oxford Univ. Press, New York, 2002).
7. E. H. Lennenberg, *Biological Foundations of Language* (Wiley, New York, 1967).
8. S. Pinker, *Language Learnability and Language Development* (Harvard Univ. Press, Cambridge, MA, 1984).
9. C. Boeckx, M. Piatelli-Palmerini, *Linguist. Rev.* **22**, 447 (2005).
10. W. T. Fitch, M. D. Hauser, *Science* **303**, 377 (2004).
11. T. Q. Gentner, K. M. Fenn, D. Margoliash, H. C. Nusbaum, *Nature* **440**, 1204 (2006).
12. R. A. Murphy, E. Mondragon, V. A. Murphy, *Science* **319**, 1849 (2008).
13. N. Chomsky, *Syntactic Structures* (Mouton, the Hague, 1957).
14. A. Friederici et al. *Proc. Nat. Acad. Sci.* **103**, 2458 (2006).
15. M. Musso et al., *Nat. Neurosci.* **6**, 774 (2003).
16. W. Enard et al., *Nature* **418**, 869 (2002).
17. S. Haesler et al., *J. Neurosci.* **24**, 3164 (2004).
18. A. R. Luria, *Traumatic Aphasia* (Mouton, the Hague, 1969).
19. J. T. Hutton, N. Arsenina, B. Kotik, A. R. Luria, *Cortex* **13**, 195 (1977).
20. T. G. Bever, C. Carrithers, W. Coward, D. J. Townsend, in *From Neurons to Reading*, A. Galaburda, Ed. (MIT Press, Cambridge, MA, 1989).
21. D. J. Townsend, C. Carrithers, T. G. Bever, *Brain Lang.* **78**, 308 (2001).
22. D. S. Ross, T. G. Bever, *Brain Lang.* **89**, 115 (2004).
23. S. Chan, thesis, University of Arizona (2007).

BIOCHEMISTRY

RT Slides Home…

Stefan G. Sarafianos[1] and Eddy Arnold[2]

To access its target sites, HIV reverse transcriptase slides and flips on nucleic acid substrates.

For HIV to replicate inside human cells, it must convert its single-stranded RNA genome into double-stranded DNA that can be integrated into the host genome (*1*). This formidable task is achieved by HIV reverse transcriptase (RT), a multifunctional enzyme that has RNA-dependent and DNA-dependent DNA polymerase activities to synthesize minus and plus DNA strands, ribonuclease H (RNase H) activity to degrade the RNA strand of the RNA-DNA replication intermediate, a strand displacement activity to remove the remaining RNA and DNA fragments to allow synthesis of the plus DNA strand, and a strand transfer activity to move newly synthesized DNA within or between templates. Although 20 years of crystallographic and biochemical studies have illuminated the molecular details of the chemistry of DNA synthesis, there have been relatively few insights into how RT finds the end of the nucleic acid substrate where it begins DNA synthesis, how it displaces nucleic acid fragments, or where and how it executes masterful leaps when transferring DNA between templates. On page 1092 of this issue, Liu *et al.* (*2*) describe elegant single-molecule fluorescence resonance energy transfer (FRET) experiments that provide a view of RT at work. They show that RT has a remarkable ability to slide on nucleic acid duplexes, rapidly shuttling between the two ends and flipping into the polymerase-competent binding mode when needed.

Important structural features of RT (*3*, *4*) and its molecular interactions with substrates and inhibitors have been elucidated through extensive crystallographic studies (*4–7*). HIV RT is an asymmetric heterodimer composed of p66 and p51 subunits that have identical amino termini. The p66 subunit has enzymatic activity, containing the spatially distinct polymerase and RNase H active sites, whereas the smaller p51 subunit plays a structural role. The p66 polymerase domain comprises four subdomains: fingers, palm, thumb, and connection. Although p51 folds into the same subdomains as the polymerase domain of p66, the positions of the subdomains relative to each other are different in p66 and p51.

In this study, Liu *et al.* use a single-molecule FRET assay to measure the position and orientation of RT relative to its nucleic acid substrate. They immobilized nucleic acid labeled at one end of the template or primer strand with the FRET acceptor fluorophore, Cy5, and immersed it in a solution containing RT molecules labeled with a FRET donor dye, Cy3, attached either at the RNase H domain or at the fingers domain of the p66 subunit. By monitoring the FRET efficiency, they were able to determine the enzyme's position on the nucleic acid substrate during each binding event. The same team (groups of Zhuang and Le Grice) recently used this approach to show that RT can rapidly switch between two orientations when it binds duplexes containing the unique polypurine RNA sequences that are primers for plus-strand synthesis (*8*). Now they show that the enzyme can slide between opposite termini on long duplexes and that the flipping and sliding kinetics are altered in the presence of nevirapine, a non-nucleoside inhibitor of HIV RT (NNRTI).

Here, the authors pose the question: How does RT efficiently locate the 3′ terminus of nascent DNA on a long duplex substrate so that it can extend it? This question is particularly important because HIV RT has relatively low processivity and must frequently locate the polymerization site after dissociation. Also, RT cleaves RNA-DNA hybrids at many different sites, and it is not well understood how it accesses these sites (*9*, *10*).

In answer to these questions, Liu *et al.* initially showed that RT binds an oligonucleotide that is the same size as its nucleic acid binding cleft (19 base pairs) only in the configuration that places its polymerase site at the 3′ end of the primer ("front-end" binding). However, when RT binds longer RNA-DNA (or DNA-DNA) substrates (38 or 56 base pairs), there is an equilibrium between front-end and back-end binding that favors front-end binding (see the figure). Therefore, the enzyme can stably bind either to the front end of the hybrid, poised for DNA extension, or to the back end, placing the RNase H domain close to the 3′ of the RNA (or DNA) template.

By following changes in FRET over time, Liu *et al.* were able to detect repeated transitions between front- and back-end bound states within a single binding event, suggesting that shuttling can occur between these

[1]Christopher S. Bond Life Sciences Center, Department of Molecular Microbiology and Immunology, University of Missouri, 1201 Rollins Street, Columbia, MO 65211, USA. [2]Center for Advanced Biotechnology and Medicine, Department of Chemistry and Chemical Biology, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854, USA. E-mail: sarafianoss@missouri.edu; arnold@cabm.rutgers.edu