

Linguistics and Learning to Read*

David Pesetsky (MIT)

For centuries, linguists have been examining how languages are put together. This investigation is possible because human languages are overwhelmingly orderly and law-governed. But the investigation is also *exciting*, because of a remarkable fact emerging from recent work: though languages differ in many ways, they are all cast from a common mold -- a "master plan" rooted in human biology. Linguists interested in this discovery try to determine exactly what this master plan is, and how it is reflected in the individual languages of the world. This work is important for many reasons. One reason is purely intellectual. By discovering the nature of human language, we arrive at a better understanding of ourselves. But another reason, less appreciated by most linguists, is social and political. Accurate understanding of human language is important when society is faced with decisions that have a linguistic basis -- for example: how we should teach reading to young children.

One of the most influential schools of thought, Whole Language, has come under severe criticism from many perspectives. Despite this, the ideas in Whole Language remain extremely popular among educators. They are passionately defended against criticism -- much of the passion being devoted to insulating the ideas against disconfirmation. Cognitive science that challenges the Whole Language view of skilled reading is dismissed (Shannon 1994) because it relies on laboratory experimentation. Standardized test results that fail to show the superiority of Whole Language classrooms are dismissed because of supposed inadequacies of the tests (Weaver et al. 1996)¹. Parental dissatisfaction is attributed to poor parent-teacher communication (Routman 1996, 64). Other dissatisfaction is blamed on right-wing propaganda (Weaver 1994, 294 ff), the influence of media (Routman 1996, 10-11), or inadequate public relations by Whole Language proponents (Routman 1996, 13-16).

Why is Whole Language defended so vigorously against counter-evidence that (as a matter of logic) should refute or reshape some of its central ideas? While there are presumably many answers to this question (including the many positive ideas packaged with the more controversial proposals), I want to focus on the factor most relevant to this workshop: the linguistics of Whole Language. Whole Language is deeply rooted in a set of compelling ideas about language and communication. If these ideas were true, they might supply an argument for Whole Language so strong that one *should* seek alternative explanations for all the counterevidence. Linguistic and psycholinguistic research, however suggests that these ideas are false.

I will devote these remarks to developing this point further. By presenting some of the background that linguists and psycholinguists bring to a consideration of Whole Language, I hope to clarify the reasons why the Massachusetts letter was written. I should emphasize that this document presents a personal view.

The Master Plan for Language

The first fact that strikes one about human language is its diversity. Languages differ greatly in such fundamentals as sound systems and word order. For example, Russian forbids "voiced" consonants (for example: d, g, b) at the end of its words, replacing them with unvoiced consonants (t, k, p). English is freer, and allows both types of consonants at the ends of its words (compare *bad* and *bat*, which would come out the same in Russian). On the other hand, Russian allows combinations of consonant sounds (*rta*, *Mstislav*, *gosudarstf*) which English forbids. The experience of growing up speaking Russian or growing up speaking English colors our linguistic habits. The Russian speaking English finds it hard to produce "b" at the end of an English word, just as the American speaking Russian finds it hard to say *Mstislav*.

Likewise, English and Japanese sentence structure appears quite different. For one thing, the verb precedes its direct object in English, but follows it in Japanese. *John met Mary* comes out as "John Mary met" in Japanese. Relative clauses follow nouns in English, but precede them in Japanese. *People who read books eat sushi* comes out in Japanese as "Book read people sushi eat". Despite this diversity,

* These remarks conform to a page limit. Consequently, the discussion of several points is compressed and incomplete.

¹Though positive results are hailed. See NCTE's reporting on Whole Language in Union City, New Jersey at <http://www.ncte.org/news/chronicle/top/wholelan.html> on the Web.

linguists have learned that languages like English, Russian and Japanese are organized in very nearly the same way. A fundamental unity underlies the appearance of diversity.

For example, all languages have rules that restrict the distribution of sounds. These restrictions are organized in much the same way in different languages, even when the details differ from language to language. Likewise, though word order differs from language to language, it too follows patterns seen in languages from all corners of the globe. A linguist encountering an unfamiliar language is often able to make remarkably accurate predictions about how that language works, simply because the same properties show up in language after language.

To give just one example: what non-German student of German has not struggled with the rules of German sentence structure (mercilessly mocked by Mark Twain in a famous essay)? In subordinate clauses, the verb is at the end. In main clauses, the main verb follows the first phrase (or comes first), unless there is an auxiliary verb, in which case it's the auxiliary verb that follows the first phrase: *Yesterday has the student who the book read also a record bought*. One might imagine that this pattern is a bizarre quirk of German. But *exactly the same rule* governs the syntax of Karitiana -- a totally unrelated language of the Amazon rain forest (Storto 1996). In Karitiana, just as in German, the verb is at the end in subordinate clauses, and in second position (occasionally first position) in main clauses. Not only that, but virtually the same rule holds in Vata (Koopman 1983) -- a language of West Africa belonging to the Kru family. Clearly, the peculiarities of German word order spring from something deep in the nature of language -- from some common denominator that links the inhabitants of the Amazon rainforest with Ivory Coast villagers, and links them both with speakers of German.

The Roots of Language: Chomsky's conjecture

The obvious common denominator is biology. Over 30 years ago, Noam Chomsky (1965) suggested that the common thread linking human languages is biological. He conjectured that a strong and very specific biological predisposition allows us -- in fact compels us -- to acquire and use very specific types of linguistic systems. He noted that this view not only could explain why languages are cut from a common mold, but could also explain the rapidity and uniformity of language acquisition by children. Around the world, language is acquired by children under the most diverse circumstances. Yet language acquisition is almost always extremely rapid and successful in a way that all sorts of other learning is not.

When Chomsky began his work, these ideas were conjectures. Detailed research on common properties of human languages was only starting, and useful work on language acquisition by children had barely begun. Now, however, it is clear that Chomsky's conjecture was substantially correct. The conjecture has been confirmed not only by our ever-increasing understanding of language structure, but also (more recently) by investigations of language acquisition that point in exactly the same directions.

For example, the research program initiated by Peter Eimas in the 1970s has established that *very specific details* of linguistic sound systems are present in infants at birth or shortly after. (Good summaries of this work are Mehler & Christophe 1995 and Eimas 1996.) In some cases, the knowledge is present despite the absence of any linguistic experience that could account for the knowledge. For example, infants display knowledge of specific linguistic contrasts among sounds that are not part their mother's native language -- *l* and *r* in Japanese infants being the best-known example. As they grow older, they seem to lose access to this knowledge -- as if this aspect of language acquisition involved "unlearning" innate knowledge. Likewise, informed investigation of syntax in very young children shows an amazing amount of linguistic knowledge already in place -- even when casual inspection of the child's speech would lead one to think the child had much yet to learn.

The message of this research is loud and clear. With each new research report, it becomes harder and harder to find an age at which a child does *not* know some particular fact about language. Some aspects of this knowledge -- the facts that distinguish one language from another -- obviously cannot be completely innate. Nonetheless, such facts seem to be learned so early that it is hard not to believe that biology plays a helping role. In other cases, we *know* that biology must be playing a crucial role.

One of the most spectacular recent studies was carried out by Iverson and Goldin-Meadow (1995). This study concerned linguistic gesturing in four congenitally blind children (about 10 years old). Gesture is an integral component of oral language -- playing much the same role as intonation.

Everyone gestures naturally and unconsciously while speaking. One might be tempted to argue that we gesture because we have seen others gesture, and have learned about its communicative impact. Certain conventionalized gestures (like the "Shh!" finger-to-lips gesture) surely *are* learned in this manner. What about the less conventionalized gestures? Do we use gesture because we've seen others do it, or do we gesture because something in our biology makes us do it? The behavior of the congenitally blind allows us to study this question. They do not have access to information about other people's gestures, and have little way of gauging the communicative impact of their own gesture. In fact, the four blind children studied by Iverson and Goldin-Meadow did *not* know the conventionalized English gestures (including the "Shh!" gesture). Nonetheless, their natural speech was accompanied by rich, fairly normal gesturing -- quite comparable to the gesturing of sighted counterparts. Why do they gesture? Since the answer can't be rooted in social interaction, the only possible answer lies within the children. Linguistic gesture is something that we, as humans, are innately designed to produce. (That is why their paper was called "What's communication got to do with it?") One can imagine no more perfect support for "Chomsky's conjecture".

Linguistics and Learning to Read

As news of this research filters out to the general public, it does not find a vacuum. People already have views about language, which form a kind of "common-sense linguistics". This "common-sense linguistics", not surprisingly, sees language as a creature of its primary function: communication. Furthermore, since language is useful, "common-sense linguistics" assumes that we acquire language as a deliberate, social act -- learning to use a tool in order to reap the benefits of the tool. From this chain of reasoning comes a view of language as a *human creation*, rather than a biologically rooted instinct.

Such a view is tempting in part because aspects of language *do* look like this. We *do* have the ability to manipulate language: for example, by coining new words and usages (*outsource, downsize, grow the economy*). Also, language *does* have a social and communicative function. Different speech styles are used, consciously or unconsciously, for social purposes. Sentences are structured to send signals about old and new information within narrative, to increase or decrease the vividness of depiction, and to engage or discourage a listener's interest in different parts of a message. Many linguists study these aspects of language, contributing richly to our knowledge about language and its uses.² But the existence of functionally motivated parts of language does not entail the non-existence of other aspects of language (Givón 1995, xvi). Just because *some* aspects of language are rooted in discourse and function, we cannot conclude that *all* aspects of language are rooted in discourse and function. This is where actual linguistics discussed above parts company with its "common-sense" precursor. This is where our troubles start.

If language were just a means of communication, we would expect a very blurry line between language and other means of communication. The biological view that arises from more systematic investigation, by contrast, suggests a sharper line -- a split between communicative systems rooted in our tool-making capacity and linguistic systems more deeply rooted in our biology. One place where this split is particularly obvious is the boundary between oral/sign language and written language.

Oral and signed language is universal across the species, largely invariant in basic design, and acquired in a rapid, automatic manner. Written language is quite different. Knowledge of written language is far from universal across the species. Written language, unlike spoken language, is generally disseminated through *teaching*. There is no evidence that children are born with knowledge of written symbols, as they are born with knowledge of speech sounds. There is no evidence that written forms of language are generally acquired in the rapid, automatic manner of spoken and signed language. Even in literate societies, there is variation in the ease of literacy acquisition that is unheard of in the domain

²There is a division within linguistics between "functionalists" who study language use and "formalists" who focus on matters of linguistic form. This does not, however, translate into support for Whole Language doctrine. A commitment to the study of language function does not entail a commitment to the view that instruction in *early reading* should give special primacy to function over form. The line between functionalist linguistics and Whole Language is easily drawn. Reaction from functionalists to the linguists' effort in Massachusetts (e.g. when discussed at the Milwaukee conference on Formal-Functional Approaches to Linguistics, April 1996) has quite consistently reflected such a line. Good functionalist work is *not* the "common-sense linguistics" discussed in the text, precisely because it pays attention to the *interplay* of form and function.

of spoken and signed languages. The two systems -- written and oral/sign -- have very different cognitive footprints. One looks like an *instinct*; the other looks like a *tool* (Lieberman & Lieberman 1990).

Nonetheless, prominent voices within the education community assume just the opposite. Let us examine two central theses of Whole Language:

1. Written language is a communicative tool for "making meaning" equal to oral language. "Written language shares all the characteristics of oral language except that it's visual rather than aural." (Goodman 1996, 13). Consequently it can be learned "naturally", just like oral language.
2. Skilled reading does not start with the decoding of speech from symbols, but with any one of a number of "strategies" that extract meaning from a page -- including reliance on contextual cues and reliance on pictures. Reading is a "psycholinguistic guessing game" (Goodman 1967).

Consider the first thesis. Since the properties of language, according to the common-sense view, are just the properties of a communication system, we expect any communicative tool to display similar properties. That is why evidence that oral/signed language is *special* poses a challenge. In response, the evidence is routinely denied:

"From a whole language perspective, language learning does not occur because it is innate; nor does it occur inevitably. But if language is an integral part of the functioning of a community and is therefore used as a social resource around and with neophytes, then the learning is *natural* in that it is 'incidental' to what else is going on. The social functions of language, not language itself, are the focus. Indeed, it is because people are profoundly social creatures and because language enables them to enact their social nature that it is learned in the first place." [Edelsky et al. 1991, 16]

"Language is not innate, and not learned as imitation...Language learning is a process of social and personal invention...Form follows function in language development." [Goodman 1987, 18]

Such denials defend Whole Language against the possibility that "natural" acquisition might be limited to oral and signed language. Whole Language assumes exactly the opposite:

"Just as they learn the patterns of oral language, so most children will unconsciously learn common phonics patterns, given ample opportunity to read environmental print and predictable and enjoyable materials, and ample opportunity to write with invented (constructive) spelling..." [Weaver 1994, 197]

"There is evidence which indicates literacy can develop in the same "natural" way as spoken language when the conditions for learning are comparable". [Newman 1985, 60]

"We now know that learning to read and learning to write are a lot like learning to talk. We would think it funny if parents hovered over their newborn's crib, chanting the sounds of language one at a time. Parents are not trying to teach language, but rather trying to communicate with their child. They do not teach children individual sounds, but instead, use and share language naturally as a part of everyday experiences. They respect and accept their baby's babblings as talk. Although different from that of grown-ups, the child's language is celebrated and accepted without criticism. It is through constant interaction with family and friends--through using language and hearing others use it in everyday situations--that children learn to talk. Our research has indicated that the same is true of learning to read and write." [NCTE, *Elementary School Practices*]

But as we've seen, the facts paint a different picture. Parents don't need to "chant the sounds of language" to their newborns because their newborns *were born knowing them!* There is no comparable evidence for infant knowledge of letters. The differences are striking. Compare the infant who easily distinguishes the *sounds* /ba/ and /pa/ with the four-year old who mixes up the *letters* **b** and **p**.³

³Whole Language also borrows from "common-sense linguistics" the assumption that language acquisition is a deliberate, social act. "[B]abies learn language for something else -- for getting or refusing a piece of apple, for investigating how the TV works..." (Edelsky et al. 1991). This is not a credible explanation for the facts discussed here, since the linguistic knowledge of the infant and young child vastly outstrips the child's ability to display this knowledge to the innocent observer. (For readable recent surveys, see Pinker 1994 and Jackendoff 1994.)

Now consider the second thesis: that skilled reading is a "psycholinguistic guessing game". What links this thesis to the first? As far as I can tell, the two theses are linked precisely by the "common-sense linguistics" described above. If the common-sense view of language were correct, our use of written language would reflect its status as a communicative system. If so, a reading teacher should encourage all reading-like activities that create communication. This is why the Whole Language teacher encourages the reader to go outside the words to "make meaning", rather than decoding the words on the page, then interpreting them. Here too there is a conflict with facts as much of the reading research community sees them, but I will leave that issue to others. I focus instead on the origin of both theses in the "common-sense linguistics" of Whole Language.

Crucially, this Whole Language view of language and language acquisition can be maintained only if one disregards the fields that study these topics. In fact, although many Whole Language texts discuss language acquisition, I have yet to find a mention of even the most classic results -- for example, the work on infant speech perception.⁴ The gap is even more striking when we observe how perfectly these results dovetail with work on the teaching of reading. Consider, for example, Foorman et al. (1996), who report a beneficial effect of systematicity -- or Pressley et al. (1996), who suggest a role for teacher-directed "scaffolding". Their conclusions conflict with the Whole Language view (since oral and sign language acquisition requires neither systematicity nor scaffolding), but make perfect sense once one recognizes the special status of oral and sign language. The same is true of the now well-established finding that "phonemic awareness" is a precondition for skilled reading (cf. Gough, Ehri and Treiman 1992). Linguistic research does not in and of itself entail these results (an important and frequently misunderstood point). But it leads us to expect strong differences between language acquisition and learning to read, and allows us to make sense of such differences when they are detected.

The Massachusetts Letter

Now, suppose you are a linguist, psycholinguist or language acquisition researcher living in Massachusetts. Your state produces a draft of a Curriculum Framework on "English/Language Arts" that fully reflects the "common wisdom" associated with Whole Language. The framework is mandated by a state "Education Reform" law, and will be used as a basis for testing and evaluation of students and schools. The document mirrors in every way the "common wisdom" of Whole Language. It presents folk theories of language acquisition as fact, stresses non-existent similarities between learning to read and language acquisition -- and even describes a putative history of recent linguistics that brings it in line with Whole Language:

"In the past, research focused on the components of language -- phonological and grammatical units. As a result, we understood and taught the language processes as separate entities characterized by discrete skills. More recently, language researchers have shifted their focus to study language from the perspective of its primary function -- communication." [Mass. Dept. of Education, 1995.]

⁴There is some debate within the acquisition literature about the "domain-specificity" of many aspects of language acquisition. This debate, due largely to developments in "connectionism", challenges one of the classic arguments for the "innatist" position -- the "poverty of the stimulus". Sophisticated learning systems can extract patterns from noisy data. Researchers wonder whether some of the miracle of language acquisition might reflect this sort of pattern extraction, rather than innate, domain-specific knowledge. This debate may ultimately bear on the questions discussed here, but leaves the facts cited above untouched. We would still need an explanation in terms of mental architecture for universal properties of language. We would also need explanations for the grossly different patterning of language acquisition and learning to read.

What should you do? Your state is set to recommend a policy that will affect children, which is based on a linguistics that you believe to be inaccurate. There are really only two choices. Do nothing (shaking your head sadly, perhaps), or join your colleagues in writing a letter. In Massachusetts, we wrote a letter.⁵ As it turned out, others within the Dept. of Education shared our concerns. Consequently, the letter -- after some initial bumps -- has contributed positively to the continuing development of the Curriculum Framework. Other people have distributed, redistributed and interpreted the letter in all sorts of strange, surprising and sometimes disturbing ways, but the letter remains what it was intended to be: a comment on a Massachusetts document which was at odds with the sciences to which we have devoted ourselves.

References

- Chomsky, Noam (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge.
- Edelsky, Carole, Bess Altwerger and Barbara Flores (1991) *Whole Language: What's the Difference?* Heinemann, Portsmouth.
- Eimas, Peter (1996) "The Perception and Representation of Speech by Infants", in James L. Morgan and Katherine Demuth (eds) *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, Laurence Erlbaum Associates, Mahwah, New Jersey.
- Foorman, Barbara, David J. Francis, Teri Beeler & Debbie Winikates, and Jack M. Fletcher (1996) "Early Interventions for Children with Reading Problems: Study Designs and Preliminary Findings", to appear in *Learning Disabilities: A Multi-Disciplinary Journal*.
- Givón, Talmy (1995) *Functionalism and Grammar*, Benjamins, Amsterdam.
- Goodman, Ken (1967) "Reading: a Psycholinguistic Guessing Game", in *Journal of the Reading Specialist* 6, 126-135.
- Goodman, Ken (1986) *What's Whole in Whole Language?* Heinemann, Portsmouth.
- Goodman, Ken (1996) *On Reading*. Heinemann, Portsmouth.
- Gough, Philip B., Linnea C. Ehri and Rebecca Treiman, eds. (1992) *Reading Acquisition*, Laurence Erlbaum Associates, Hillsdale, New Jersey.
- Iverson, Jana and Susan Goldin-Meadow (1995) "What's Communication Got to Do with It?", paper presented at the 1995 Boston University Conference on Language Development; unpublished University of Chicago ms.
- Jackendoff, Ray (1994) *Patterns in the Mind*, Basic Books. [also in paperback]
- Koopman, Hilda (1983) *The Syntax of Verbs: from Verb Movement Rules in the Kru Language to Universal Grammar*, Foris Publications, Dordrecht.
- Lieberman, I.Y. and A.M. Lieberman (1990), "Whole Language versus Code Emphasis: Underlying Assumptions and their Implications for Reading Instruction", *Annals of Dyslexia* 40, 51-76. Reprinted in Gough et al. (1992).
- Massachusetts Department of Education (1995), *Conveying and Constructing Meaning*, first draft.
- Mehler, Jacques and Anne Christophe, "Maturation and Learning of Language in the First Year of Life", in Michael Gazzaniga (ed.) *The Cognitive Neurosciences*, MIT Press (1995)
- Newman, Judith (1985) *Whole Language: Theory in Use*. Heinemann. Portsmouth.
- National Council of Teachers of English (1995?), *Elementary School Practices*, brochure for parents. (<http://www.ncte.org/idea/language/elem.html>)
- Pinker, Steven (1994) *The Language Instinct*, Morrow. (also in paperback)

⁵The letter was sent to the Commissioner of Education, not to the legislature, as claimed in the Workshop prospectus. We were also not responsible for the circulation of this letter in Ohio, as the same prospectus implies.

Pressley, Michael, Ruth Wharton-McDonald and Jennifer Mistretta (1996) "Effective Beginning Literacy Instruction: Dialectical, Scaffolded, and Contextualized", SUNY Albany ms.

Routman, Regie (1996) *Literacy at the Crossroads*. Heinemann, Portsmouth

Shannon, Patrick (1994) "People Who Live in Glass Houses", in Carl B. Smith (ed.) *Whole Language: The Debate*, EdInfo Press.

Storto, Luciana (1996) "Verb Raising and Word Order Variation in Karitiana", unpublished ms., MIT.

Weaver, Constance (1994) *Reading Process and Practice*, Heinemann, Portsmouth.

Weaver, Constance, Lorraine Gillmeister-Krause and Grace Vento-Zogby (1996) *Creating Support for Effective Literacy Education: Workshop Materials and Handouts* [fact sheet: "Facts on Myths about Whole Language Education" from the Web at <http://www.heinemann.com/hbbc/08894f7.html>]