

In: Mattingly, I.G. and M. Studdert-Kennedy (eds.), *Modularity and The Motor Theory of Speech Perception* (1991). 339-346

Reading and the Biological Function of Linguistic Representations

Ignatius G. Mattingly
Haskins Laboratories

Abstract

The processes of reading and writing present a problem for those who claim that linguistic processing is modular. How is it that the language module, specialized to respond only to speech-like acoustic patterns, can apparently respond also to optical patterns of arbitrary form? It is proposed that cognitive linguistic representations such as those that the module creates can also drive the module, even if these representations are incomplete, and even if they are actually of cognitive origin. Thus, given a convention for transcribing such incomplete representations (an orthography), it is possible to exploit the language module in reading and writing. But what is the biological function of an arrangement in which cognitive linguistic representations and not just semantic ones can be modular inputs as well as modular outputs? Not communication, which could have been managed more straightforwardly without linguistic representations external to the module. The function of this arrangement is rather to bring about language change and diversity and thus, as Nottebohm has proposed, to facilitate subspeciation.

Lieberman and I (1985) have embraced Fodor's (1983) proposal that language is the business of a module. We have also, however, repeatedly urged that the "secondary" processes of reading and writing are closely related to and dependent on the "primary" processes of speaking and understanding speech (e.g., Mattingly, 1972; Lieberman, Lieberman, Mattingly, & Shankweiler, 1980). It would seem to follow that reading and writing must somehow make use of the language module. In some ways, this seems quite plausible: Many of the properties that, according to Fodor (1983, Part III), characterize the modular input

processes generally and the understanding of spoken language in particular hold also for reading, at least (and this is obviously a significant reservation) in the case of the mature, experienced reader. Thus, reading is "mandatory" (cf. III. 2); one can't look at a word without reading it. "Central access" is limited (cf. III. 3): One has no intuitions about the module's computations and internal representations. Reading is fast (cf. III. 4)—even faster than listening. Reading is "informationally encapsulated" from "top-down" cognitive influences (cf. III. 5), at least to the same extent as understanding speech. If reading has these modular properties, and if we exclude the possibility that there is a separate reading module, it must be that the language module is used in reading.

There is, however, a difficulty. A module is supposed to be "domain specific" (III. 1), and as long as attention is restricted to the understanding of speech, the language module appears to meet this requirement: Its domain is the domain of phonetic events, clearly a significant natural class, and the only acoustic signals to which it attributes linguistic structure are those that are (or could have been) produced by human vocal tract movements. But if we are going to claim that the linguistic module operates in reading, then we apparently have to say that the module can also attribute linguistic structure to signals that result from various sets of arbitrary optical patterns: the signs of writing systems. (A superficially similar problem is posed by manual sign language: see, in this volume: Poizner, Klima, & Bellugi, chapter 7; Neville, chapter 11; and Studdert-Kennedy, chapter 4 comments). These written signs (unlike manual signs) do not inherently constitute a natural class of any interesting kind, and if they specify phonetic events, it is only by convention.

I believe that there is a way out of this difficulty and that we can continue to view language as modular, but it will be necessary to elaborate the account of the language module presented in Fodor (1983). I hope to do so in a way that cannot be dismissed as *ad hoc*. The revised language module that results, however, seems at first sight rather peculiar. It therefore becomes necessary to explain why this kind of language module is reasonable from an evolutionary viewpoint.

According to Fodor's account, the language module is an "input system." When an utterance is understood, the module provides to central cognition a representation of its linguistic form—that is, its phonetic, phonological, morphological, and syntactic structure. It is also suggested that "logical form" is represented, but whether (and if so, how) the speaker's "message" (in the sense of Fodor, Bever, & Garrett, 1974) is represented or centrally derived from logical form is problematic.

The first way in which I want to elaborate Fodor's account is to regard the production of utterances, as well as their perception, as modular. I'm going to assume also that one module is responsible for both processes, and not two distinct modules, but this is really a separate issue (see Liberman & Mattingly, 1985; chapter 18, this volume). This assumption requires that the language module be able to operate in different modes, the input in one mode perhaps

being the output in another, although the same internal computations go on in all modes. As will be seen, this assumption is required on other grounds, anyway.

The input to the module, in its production mode, is the speaker's message (or its logical form). The output is the neuromotor commands to the speech articulators. But what is of present interest is that there is a second output, a linguistic representation of the utterance that is cognitively accessible to the speaker. Although he did not choose his words (the module did), he knows what he said as well as what he meant.

A further elaboration is that although in perception and production linguistic representations are outputs of the module, they can also be inputs. This is most obvious in the case of short-term rehearsal. The linguistic representation of an utterance just heard or spoken decays; in order to retain what has been said verbatim, one requires the module to recompute the utterance, thus creating a fresh linguistic representation. (Thus, for the modularist, linguistic processes support short-term memory processes and not the other way around; see Ren & Mattingly, 1989.) The input to this computation cannot be the original signal, which is long since lost; nor is it the message, for such an input would lead to a series of increasingly inaccurate paraphrases as rehearsal continued; it is the decaying linguistic representation itself. Because this representation is highly redundant, it can be accurately refreshed by the language module.

Next, and crucially for my account of reading, I claim that linguistic representations sufficient to drive the language module need not themselves have been produced by the module; they may be synthetic, that is, of purely central origin. A speaker-hearer can decide to think of words that begin with *b*, or words that rhyme with June, or he can rearrange a sentence in the active voice as a sentence in passive voice. That linguistic representations can be thus centrally devised and manipulated is hardly surprising; they are mental objects and can be freely analyzed into their component parts and these parts recombined. Inventing pseudowords is no different, in principle, from inventing centaurs and griffins. What is significant is that these purely synthetic representations can then be modular inputs: We hear the corresponding utterances in the mind's ear. Indeed, it is impossible to synthesize a word without so hearing it: The mandatoriness of the module is still evident when the input is synthetic.

The kinds of mental acts I have just been talking about have sometimes been called "metalinguistic." From the modularist's point of view, however, this term is misleading. When I think of some phonological sequence and use the language module to compute its full linguistic representation, this is still natural, first-order linguistic behavior. But there is nevertheless an important difference between such mental acts and producing and understanding speech: Awareness of the structure of linguistic representations (Mattingly, 1972) is needed. Cognitive representation in itself does not guarantee such awareness; it merely makes it possible. Linguistic representations are quite complex, and one has to learn to attend to particular aspects of these representations, just as the painter learns to

attend to details of visual representations. On the other hand, linguistic awareness does not imply access to representations internal to the module.

Synthetic linguistic representations are usually incomplete; it would take a lot of mental effort to synthesize a complete one. But this does not prevent them from serving as inputs to the module. The module seems quite happy to produce a complete linguistic representation that will consist with whatever sort of partial representation the speaker-hearer has synthesized. For example, the speaker-hearer can compose candidate sentences just by mentally concatenating arbitrarily chosen words represented as strings of segmental phonemes. If a candidate happens to be actually grammatical, the phrase-structure and prosody as well as the segmental structure will be computed by the module and cognitively represented. Note also that, for a speaker, the input can be a message together with a partial linguistic representation: A speaker *can* choose his words, if he wants to. Similarly, he can formulate an utterance under systematic phonological constraints; that is, write verse.

We have now postulated several sorts of input to the language module: acoustic signals, "messages," and linguistic representations, both modular and synthetic. That the module seems able to accept so many different forms of input has implications about the sort of processing that must go on inside it. Some form of analysis-by-synthesis, in which any of several forms of input may serve as a filter specification, may be indicated. However, I will not pursue this issue here.

It should now be pretty obvious how my account of writing and reading goes. The writer uses the language module to produce an utterance corresponding to his message. This yields a linguistic representation. Because he has mastered a system for partially transcribing linguistic representations—that is, an orthography—he can write down the text of the utterance. The reader, who also knows the orthography, uses the text to synthesize a partial linguistic representation. From this, the module computes a complete linguistic representation, plus logical form, and the text is understood.

This account of reading and writing is not very different from others that have been presented in nonmodular terms; it makes these activities dependent on the mechanisms for spoken language, but it recognizes that written symbols and speech signals stand in a quite different relation to the linguistic structure of the utterance and are differently processed. The special role of central processes in reading and writing is accommodated. The paradox I began with is avoided, for it is not the arbitrary orthographic symbols themselves that turn the language module on but rather the linguistic representations cognitively derived from these symbols.

It has, however, been necessary to complicate the language module considerably. It no longer looks like your garden-variety input system (stereopsis, scene analysis, echo ranging), having acquired various alternative modes and inputs that it did not have in Fodor's original version. I have tried to show that these elaborations could be justified independently of the requirements of reading and writing, but this may just make things worse for the devout modularist.

Before trying to sort this out, I want to indicate some other little puzzles that arise even with the basic Fodor version of the language module. First, why is there a linguistic representation at all? Only logical form is directly pertinent to the speaker's message. But the listener is nevertheless presented by the module with a wealth of syntactic, morphological, phonological, and phonetic information. Compare the actual language module with a hypothetical one that would encode from mentalese to neuromotor commands and decode from the acoustic signal to mentalese, with no linguistic representations external to the module. Why must the language module trouble us with all these linguistic details?

Note that this is not a problem for nonmodular accounts, which regard this "extra" information as a byproduct of perceptual processes. Thus, phonetic representations are supposed to be there, because they were used by working memory to support the parser's computations (Baddeley & Hitch, 1974). But the modularist cannot resort to this sort of explanation, for intermediate representations, internal to a module, are supposedly inaccessible. Therefore, all externally available representations that the module produces are true outputs and have to be explained and biologically justified as such.

Secondly, there is the redundancy of the linguistic representation. It includes, for example, both a systematic phonemic level and a phonetic level. Yet the latter is more or less predictable from the former. That both should be computed is unsurprising, but why should both be centrally represented?

Finally, there is the problem of language change. Why should there be language change and thus linguistic diversity? It is often thought that language change results from drift, or noise in the system. However, this explanation seems inconsistent with the modularity of language. Central processes are supposed to be sloppy, imprecise, imperfect, but not modular ones. If a stable scene analysis module could evolve, why not a stable linguistic module?

I suggest that these puzzles can be satisfactorily answered only if we revise our ideas about the biological function of language. We are apt to suppose that this function is primarily the communication of messages. Is it not the superior ability of our species to communicate that put it at the top of the tree? If only something could be done about linguistic diversity, which can interfere seriously with communication, the situation would be ideal.

There is, however, another possible account, according to which communication of messages is just the come-on. On this account, the true function of language is to divide the species into isolated subpopulations. So much, indeed, is implied by the myth of Babel. Nottebohm (1970), however, made the idea more precise: Genetically isolated subpopulations can adapt rapidly to local environmental conditions. Dialects, of bird song or of language, facilitate such isolation. However, because the dialects are learned, the isolation can be accomplished without irrevocable commitment to actual subspeciation.

We can now begin to understand some of the peculiarities of the language module and, in particular, the function of the linguistic representation. It is there, in the first place, to tell the listener what dialect group the speaker belongs to.

This is the information that is really important to communicate. It is of some interest to note how the module does this. Given input in a dialect different from the listener's, the language module does not simply halt (another fact with implications about linguistic processing that I will not pursue here). It is, as they say, "user-friendly." It makes the best analysis it can and at the same time indicates discrepancies between phonological and phonetic levels that the listener is instantly aware of, as if to say: "This word is doubtless supposed to be /pil/, but the [p] is not sufficiently aspirated." Of course, the more different the speaker's dialect is from the listener's, the less coherent the module's analysis will be, but even under conditions of total mutual unintelligibility, the module provides *some* pattern of discrepancies, and this pattern is what specifies the speaker's dialect for the listener. The point is that the method of indicating the discrepancies depends on their being multiple levels in the linguistic representation. These levels are redundant from the standpoint of grammatical description, but not in the actual ecological situation, in which the dialects of speaker and listener cannot be assumed to be identical.

The linguistic representation, therefore, gives the listener information about the dialect of the speaker. But at this point, some one might raise the objection that there is still no reason for this information to be centrally represented. Consider another hypothetical language module, in which a sufficient amount of dialect discrepancy leads automatically and subdoxastically to an aversive reaction by the listener without central access to the linguistic details. But, though things might work in just this way for white-crowned sparrows, the case is different with human beings. It may be true that "Every time an Englishman opens his mouth, he makes some other Englishman despise him," but this is not the inevitable result of an automatic, mandatory, domain-specific, subdoxastic, and so forth, process. It depends rather on the political and social beliefs and attitudes of the listener. Having perceived that the speaker's dialect is different from his own, the listener might indeed decide to avoid the speaker or to attack him. On the other hand, if the listener regards the speaker's dialect group as prestigious, he may make quite a different decision: to imitate or adopt the speaker's dialect.

Whether the difference between the dialects of speaker and listener is large or small, this decision of the listener's entails some form of second-language acquisition, and, as a consequence, the linguistic representation comes into play in another way. Both first-language acquisition and second-language acquisition entail modification of the language module. But one difference between the two is that whereas first-language acquisition depends on the acoustic input during the critical period and is mandatory, second-language acquisition (setting aside "infant bilingualism") occurs only when the speaker-hearer has the appropriate beliefs and attitudes, and requires central synthesis of linguistic representations to drive the language module. That is, practice is required, not merely immersion. The language learner attends to his linguistic representations of native

speaker's utterances, synthesizes similar representations of his own, produces the corresponding utterances, and compares the resulting output representation with the native model. From the module's point of view, the synthetic input is discrepant, in the sense suggested earlier. The only way for the module to make sense of it is to revise its internal algorithms. Second-language learning can be viewed as a series of such forced revisions.

It has often been suggested that bilingual speakers play a crucial role in changes that result in new dialects and languages. Having learned a second language, L2, their first language, L1, is to some degree altered, or perhaps we should say that the algorithms of their language modules are partly consistent and partly inconsistent with both L1 and L2, and they really specify a third language, L3. If there are a large number of such speakers, the way is open for language change, for the next generation of speakers will be acquiring a version of L3 rather than of L1.

Linguistic representations thus have two functions in the system for genetic isolation that the language module supports. On the one hand, they allow individuals to identify the subgroups to which other individuals belong. On the other hand, by virtue of their role in second-language acquisition, they facilitate the emergence of new dialects and languages in response to changing social and political conditions.

Viewed primarily as a genetic isolation system, rather than primarily as a system for communicating messages, the revised language module appears somewhat less bizarre. We have a way to think about certain linguistic activities—reading and writing and also second-language acquisition—in which central processes clearly have an important role. Furthermore, although it has acquired some more inputs, the language module still retains its original endearing properties. It is still domain-specific, fast, mandatory, and so on, and although central processes can provide input, they cannot tamper with the processing of such input by the module. In this sense, the criterion of “limited central access” is still satisfied.

Acknowledgment

Support from NICHD grant 01994 to Haskins Laboratories is gratefully acknowledged.

References

- Baddeley, A. D., & Hitch, G. B. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and activation* (vol. 4, pp. 47–90). New York: Academic Press.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

- Fodor, J. A., Bever, T. M., & Garrett, M. (1974). *The psychology of language*. Cambridge, MA: MIT Press.
- Liberman, A. M., & Mattingly, I. G. (1985). The Motor Theory of Speech Perception revised. *Cognition*, 21, 1-36.
- Liberman, I. Y., Liberman, A. M., Mattingly, I. G., & Shankweiler, D. (1980). Orthography and the beginning reader. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading, & dyslexia* (pp. 137-153). Baltimore: University Park Press.
- Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading* (pp. 133-147). Cambridge, MA: MIT Press.
- Nottebohm, F. (1970). Ontogeny of bird song. *Science*, 167, 950-956.
- Ren, N., & Mattingly, I. G. (1989). *Short-term serial recall performance by good and poor readers of Chinese*. Manuscript submitted for publication.