Simply imagine that someone who understands no Chinese is locked in a room with a lot of Chinese symbols and a computer program for answering questions in Chinese. The input to the system consists in Chinese symbols in the form of questions; the output of the system consists in Chinese symbols in answer to the questions. We might suppose that the program is so good that the answers to the questions are indistinguishable from those of a native Chinese speaker. But all the same, neither the person inside nor any other part of the system literally understands Chinese...


In 1972 John Searle wrote an article in The New York Review of Books called “Chomsky’s Revolution in Linguistics” (Searle 1972). This article did much to explain to a non-specialist audience what was revolutionary about generative grammar, and the implications of the new theory for our understanding of the nature of language and mind. The first part of this article remains the best short summary of the “Chomsky revolution” I have seen.

It is thus disconcerting that, thirty years later, Searle has returned to The New York Review with “End of the Revolution,” or, as the cover has it, “Chomsky’s Thermidor.” Searle writes that “the revolution has not succeeded. Something else may have succeeded, or may eventually succeed, but the goals of the original revolution have been altered and in a sense abandoned.” This announcement may come as something of a surprise to many generative linguists, who are under the impression that they are still advancing the general program initiated by Chomsky.

How did the end come and why were we not told? Searle (1972) sums up the revolution in a chart:

<table>
<thead>
<tr>
<th>Structuralism</th>
<th>Generative Grammar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Matter</td>
<td>corpus of utterances</td>
</tr>
<tr>
<td></td>
<td>speaker’s knowledge of how to produce and understand</td>
</tr>
<tr>
<td></td>
<td>sentences, his linguistic competence</td>
</tr>
<tr>
<td>Goal</td>
<td>classification of the elements of the corpus</td>
</tr>
<tr>
<td></td>
<td>specification of the grammatical rules underlying the</td>
</tr>
<tr>
<td></td>
<td>construction of sentences</td>
</tr>
<tr>
<td>Methods</td>
<td>discovery procedures</td>
</tr>
<tr>
<td></td>
<td>evaluation procedures</td>
</tr>
</tbody>
</table>

Searle (1972) shows how the shift from structuralism (in particular, the version prominent in American linguistics) to generative grammar involved a shift from an empiricist to a rationalist orientation in all aspects of the field, including its subject matter, goals, and methods. American structuralists had maintained that a language, for purposes of scientific investigation, inheres in a “corpus,” some (necessarily finite) record of a sample of the language being studied. Chomsky argued that the real subject matter of linguistic theory must be, and indeed has always been, the knowledge that speakers have of their language.

With respect to methods, American structuralists aimed to develop a set of “discovery procedures” that could automatically extract the phonemes, morphemes, and other classifiable
linguistic units from a corpus. Chomsky maintained that a grammar is in effect a theory of speakers’ knowledge of language. Since no other science has achieved, or even seeks, a mechanical procedure for discovering the true theory, the quest for discovery procedures was a hopeless enterprise. Chomsky proposed instead that linguists, like other scientists, are under no obligation to explain how they arrive at their hypotheses. The best they can do is to develop criteria for evaluating competing hypotheses.

Finally, American structuralists defined linguistics as a taxonomic science whose goal was to classify the various elements of the corpus. However, if a language is not identified with a corpus but is rather a system of knowledge, then a more ambitious goal suggests itself, and that is to specify precisely what this knowledge is. To account for the generative capacity of language, this knowledge must include some system of rules for producing and understanding novel utterances.

No matter what version of generative grammar they adhere to, almost all generative linguists looking at Searle’s chart today would maintain that nothing essential in it has changed. One might be struck by a conspicuous omission in the chart, however. Generative grammar has not one goal, but two. The goal listed in Searle’s chart is the attainment of what Chomsky has called “descriptive adequacy,” the specification of the correct grammars (systems of knowledge) for particular languages. The program of generative linguistics is equally committed to achieving “explanatory adequacy,” explaining how learners arrive at the correct grammar upon exposure to data. To achieve this level of explanation we must go beyond the specification of particular grammars and develop a theory of the cognitive principles that learners bring to the task of language acquisition. These (presumably innate) cognitive principles are known in the field as Universal Grammar (UG). Thus, the second goal of generative grammar is to arrive at a theory of UG that will project a grammar from the data available to the learner.

With the benefit of hindsight, we can perceive that though Searle 1972 appears to present a fairly orthodox account of the “Chomsky revolution,” Searle’s understanding of what this involved was already radically different from that of most linguists. The Chomsky revolution he was writing about did indeed come to an end some time ago, maybe soon after 1972.

Searle’s thought over the years has been quite consistent, and revolves around his interpretation of the notions of “rule” and “intentionality.” For Searle, a rule is something that one actually follows, or is capable of following. Not unconsciously, or “tacitly,” as linguists like to say, but consciously, and with intentionality. A nice example is a speech act (Searle 1969): someone who performs a speech act (“I declare that this congress is now open.”) engages intentionally in a rule-governed activity. For Searle, there are no unconscious rules: what linguists call an unconscious rule is really neural hardware.

Searle (1990: 591) gives the example of looking out the window of a moving car. The scene does not jump around, despite the movement of the car, because our eyeballs move to compensate for the movement of our heads. We might say that our eyes follow an unconscious rule, but this would be wrong. “What actually happens is that fluid movements in the semicircular canals of the inner ear trigger a sequence of neuron firings that enter the brain over the eighth cranial nerve.”

Searle appears to have understood the original program of generative grammar, as embodied in the Aspects model of grammar, for example (Chomsky 1965), as specifying the intentional rules of particular languages. Generative linguists did not regard the phrase structure and transformational rules of the Aspects model in this light, and indeed they fall somewhat short of the full-fledged intentionality that Searle requires of rules. Thus, Searle (1972) argued that the conception of rules of grammar in the standard theory is too syntactic and insufficiently semantic. As the Chinese room argument shows, a purely formal syntactic algorithm cannot understand anything, nor attain intentionality. However, Searle evidently felt that, with some improvements, the original program of generative grammar was capable of becoming the right sort of theory of rule following.
“It was a beautiful theory,” he writes in 2002. “But the effort to obtain sets of such rules that could generate all and only the sentences of a natural language failed...seen from outside a striking feature of the failure is that in Chomsky’s later work even the apparently most well-substantiated rules, such as the rule for forming passive sentences from active sentences, have been quietly given up.”

Quietly? The key phrase here is “seen from outside.” The old passive transformation, as it turns out, provides an excellent illustration of the tension between descriptive and explanatory adequacy. A formulation of this rule that Searle might have in mind can be found, for example, in Burt 1971: 37:

Structural Description:  
```
## (PreS) — NP — [Aux X Passive Aux] V (Prep) — NP — Y
```

⇒ Structural Change:  
```
1 4 3 5 by 2
```

This rule raises some difficult questions with respect to explanatory adequacy. If transformational rules can introduce a by phrase, what other specific phrases can they mention? Is it a coincidence that the derived structure is similar to one that could be generated independently by phrase structure rules? How can a language learner arrive at precisely this rule from among the infinite number of similar highly specific transformations the theory makes available? What common basis might exist between the English passive rule and the rules of other languages?

Answers to these questions took a number of different directions. David Perlmutter argued persuasively that the passive relation may take many forms in different languages—some may express the passive syntactically (though not the way English does), others may do so morphologically, and so on. If the essence of the “passive” transformation is that it expresses the passive relation, then to be generalized it must refer to grammatical relations. The theory of Relational Grammar (e.g. Perlmutter 1983) developed this idea.

Chomsky and others in generative grammar sought to generalize transformational rules in another direction, by looking for universal constraints and further limiting the expressive power of transformations. This direction led away from construction-specific rules. Thus, the old passive rule was generalized and simplified to Move NP, then to Move α, and ultimately to just Move, with general conditions serving to limit the application of the rule.

Thus, the rule of passive was not quietly dropped—on the contrary, debates about how to generalize rules like passive caused a considerable ruckus in the field, though the noise may not have reached Searle’s office. Whichever direction one took, however, keeping the rule as it was was not an option if one hoped for a more explanatory theory.

From Searle’s point of view, however, cashing in construction-specific rules that are not too far from being real (intentional) rules in exchange for general rules of UG is a “stunning mistake.” For Searle, UG is neural hardware, which has nothing to do with rules. To a working linguist, substituting some unspecified neural mechanisms for explicit principles of UG is not a good bargain. Nor is the generalization of construction-specific rules to the status of principles of UG an “abandonment” of the revolution.

But Searle is not swayed by such considerations, as I have had occasion to personally discover (see also Bromberger 2002, Chomsky 2002, and Searle’s replies in Searle 2002b, 2002c). Dresher and Hornstein (1990) pleaded the case for a UG containing principles of binding and notions of anaphors and antecedents as a way of explaining why Each other like the men is ungrammatical. We observed that tossing out the formulation in (1) in favour of (2) does nothing to advance our understanding of the phenomenon, and indeed, amounts to giving up a possible explanation:
When I hear the sentence, “Each other like the men,” it strikes me as odd because I follow the deep unconscious rules [of UG specified elsewhere in the commentary].

When I hear the sentence, “Each other like the men,” it strikes me as odd because the hardware mechanism of the brain produces this effect in me.

It can be added that Searle’s examples from nonlinguistic domains do not help his case with respect to UG. For until such time as neural researchers discover anaphora fluid sloshing around in the ear canals it appears that an unconscious, nonintentional version of UG is the only way forward.

*       *       *

Searle’s Chinese Room argument is one of the most famous arguments in the philosophy of cognitive science, and has given rise to many spin-offs. Here is a variant I call the “Linguistics Room”:

* Imagine that someone who knows no linguistics is locked in a room with a lot of linguistic symbols and a fixed notion of what a “rule” must be. The input to the system consists in random linguistics publications at ten-year intervals; the output of the system consists in a critique of the linguistic theory embodied in the publications. We might suppose that the program is so good that the critique is indistinguishable from one that could be produced by someone who actually does linguistics. But all the same, neither the person inside nor any other part of the system literally understands linguistics...

Is this possible? Discuss.

References