

NGÔN NGỮ HỌC TẠO SINH CỦA N.CHOMSKY: MÔ HÌNH NGÔN NGỮ THỨ NHẤT

GS.TS NGUYỄN THIÊN GIÁP

Như ta đã biết, đối tượng nghiên cứu của ngôn ngữ học tạo sinh là *ngữ năng* của người nói lí tưởng*, vậy phải miêu tả ngôn ngữ như thế nào để thể hiện được ngữ năng của người nói, tức là thể hiện được khả năng tạo ra vô hạn các câu từ một số hữu hạn các cú liệu đã biết và khả năng hiểu được những câu chưa bắt gặp bao giờ. Trong cuốn *Các cấu trúc cú pháp* (1957), Chomsky đã trình bày *Mô hình ngôn ngữ thứ nhất* (The First Linguistic Model).

Mô hình ngôn ngữ thứ nhất của Chomsky là mô hình đầu tiên miêu tả hình thức hóa ngôn ngữ. Theo Chomsky, chỉ có miêu tả và phân tích hình thức hóa mới có thể làm cho các câu đơn giản, rõ ràng, có tính hồi quy. Ông phát hiện cú pháp rất giống với việc sắp xếp tổ hợp trong toán học.

Cuốn *Các cấu trúc cú pháp* được xuất bản năm 1957 hoàn toàn không ăn nhập với dòng chính thống đang thịnh hành của ngôn ngữ học, nó dường như một cuộc cách mạng. Ngay trang đầu tiên, ông đã tuyên bố rằng ngữ pháp là một hệ thống tự trị, độc lập với ngữ nghĩa học và việc nghiên cứu cách dùng ngôn ngữ trong các tình huống. Hơn nữa, nó sẽ được hình thức hóa như một hệ thống các quy tắc tạo sinh tập hợp vô hạn các câu.

Cách tiếp cận này tương phản đột ngột với ngôn ngữ học chính thống đang thịnh hành khi đó tin rằng vận dụng các thủ pháp thích hợp với vốn dữ liệu sẽ mang lại sự miêu tả ngữ pháp. Chomsky không chấp nhận sử dụng vốn dữ liệu hữu hạn để miêu tả mà đề nghị sự thỏa đáng của ngữ pháp được đánh giá bằng nó có thể tạo sinh một số vô hạn các câu đúng ngữ pháp và miêu tả các kiểu đánh giá trực giác mà người bản ngữ đã có về ngôn ngữ của họ. Thuộc những đánh giá này là: một chuỗi từ, đặc biệt là một chuỗi mới đã là hay không là một câu hợp thức; các câu nào đó là các câu lưỡng nghĩa (ambiguous), tức là chỉ một câu duy nhất có thể có nhiều hơn một cách giải thích, các câu khác biệt có thể mô phỏng lẫn nhau, tức là các câu khác biệt có thể có những sự giải thích đồng nhất; các kiểu câu nào đó (khẳng định và phủ định, trần thuật và nghi vấn,...) có thể quan hệ lẫn nhau một cách hệ thống và v.v.. Những đánh giá loại này tạo thành cái mà người nói hiểu về ngôn ngữ của họ, và thêm vào việc miêu tả tính hợp thức của các câu của ngôn ngữ, ngữ pháp cũng sẽ miêu tả tri thức đó.

.....

* Nguyễn Thiên Giáp, Ngôn ngữ học tạo sinh của N.Chomsky: đối tượng và mục đích, *T/c Ngôn ngữ*, Số 4, 2012.

Như trên đã nói, Chomsky đề nghị rằng ngữ pháp sẽ được coi như một hệ thống tự trị, độc lập với các hệ thống ngữ nghĩa học và âm vị học, mặc dù có quan hệ với chúng. Hơn nữa, ông đã đề nghị rằng bản thân cú pháp sẽ bao gồm một số bình diện khác biệt nhưng có quan hệ, mỗi một bình diện được đặc trưng bằng các kiểu quy tắc khác biệt và mỗi bình diện gánh chịu một phần riêng biệt của gánh nặng miêu tả. Hai bình diện quan trọng nhất là: *cấu trúc đoạn ngữ* và *các thành tố cải biến*.

Như ta biết ngữ pháp *cấu trúc đoạn ngữ* (phrase structure) là một kiểu ngữ pháp của cấu trúc luận Mỹ. Ngữ pháp cấu trúc đoạn ngữ miêu tả cấu trúc cú pháp của câu như là cấu trúc thành tố, tức là như tôn ti của các yếu tố có trật tự. Trong khung của ngữ pháp cải biến, *ngữ pháp cấu trúc đoạn ngữ* được giải thích như *quy tắc viết lại*, chẳng hạn: $S \rightarrow NP + VP$ tương ứng với câu bao gồm một đoạn ngữ danh từ và một đoạn ngữ vị từ.

Thành tố cấu trúc đoạn ngữ bao gồm một tập hợp các quy tắc cấu trúc đoạn ngữ, nó hình thức hóa một số cách hiểu sâu sắc của truyền thống về phân tích cấu trúc thành tố.

Chẳng hạn, hãy xem xét tập hợp các quy tắc sau đây:

Câu $\rightarrow NP + VP$

NP $\rightarrow T + N + \text{số}$

Số $\rightarrow (\text{số ít, số nhiều})$

VP $\rightarrow \text{Vị từ} + NP$

Vị từ $\rightarrow \text{trợ vị từ} + V$

Trợ vị từ $\rightarrow \text{thì}$

Thì $\rightarrow \{\text{hiện tại, quá khứ}\}$

T $\rightarrow \text{the}$

N $\rightarrow \text{man, ball, ...}$

V $\rightarrow \text{hit, took, ...}$

Mỗi quy tắc là một chỉ dẫn về sự viết lại phù hiệu ở bên trái của hàng như là phù hiệu hoặc các phù hiệu ở bên phải: Nó có thể được xây dựng như “phạm trù bên trái của hàng có (các) thành tố biểu hiện ở bên phải của hàng”. Sự phái sinh từ ngữ pháp này về sau có thể được trình bày bằng hình cây. Chomsky gọi các cấu trúc được tạo sinh bằng các quy tắc cấu trúc đoạn ngữ là *các cấu trúc cơ sở* (underlying structures).

Hãy xem xét một thí dụ đơn giản chứa đựng sự phân tích một câu đơn: *The cat sat on the mat* (*Con mèo ngồi trên thảm*).

Ngữ pháp cấu trúc đoạn ngữ miêu tả câu này là:

Cú pháp:

$S \rightarrow NP \quad VP$

$NP \rightarrow \text{quán từ } N$

$VP \rightarrow V(1) \quad PP$

$PP \rightarrow \text{Giới từ } NP$

Từ vựng:

Cat N

Mat N

On giới từ

Sat V(1)

The quán từ

Dù ngữ pháp này là đơn giản, nó đã lập thức phù hợp với một số nguyên lý chung. Nguyên lý chung nhất của nó là ngữ pháp bao gồm một số các

thành tố khác nhau, trong trường hợp này có hai thành tố là cú pháp (xác định thành tố được phép của cấu trúc) và từ vựng (liệt kê các từ trong ngôn ngữ và lớp từ vựng mà mỗi từ thuộc vào). Bản thân các quy tắc cú pháp được chế định theo các tuyến sau đây:

1. Tất cả các quy tắc đều có hình thức $A \rightarrow B C$.

2. $\rightarrow$ được thuyết giải như “có các thành tố”.

3. Quy tắc có thể chỉ bao gồm một phạm trù ở bên trái của $\rightarrow$.

4. Quy tắc có thể bao gồm một hoặc hơn một phạm trù (bao gồm cả các thí dụ chi tiết về phù hiệu đầu S) ở bên tay phải của $\rightarrow$.

5. Các phạm trù được trình bày ở bên phải của $\rightarrow$ được sắp xếp đôi với nhau.

6. S là phù hiệu đầu tiên, tức là sự phái sinh của bất cứ câu nào cũng phải bắt đầu với phù hiệu này.

7. Khi bên trái của quy tắc là một phạm trù đoạn ngữ thì bên phải của quy tắc cũng phải bao gồm phạm trù từ vựng tương ứng, chẳng hạn, một NP phải có một N với tư cách là một thành tố của nó (và có thể có các phạm trù khác (định ngữ)).

8. Các phạm trù từ vựng N, V, P, Det,... là vốn từ kết thúc (terminal vocabulary); tức là các phù hiệu này kết thúc sự phái sinh và không thể phát triển hơn trong cú pháp.

9. Các phạm trù từ vựng có thể tăng lên để chỉ thành viên của lớp nhỏ của phạm trù, chẳng hạn, phạm trù V được phân biệt thành V(1) (*lay*, *sat*), phân biệt với V(2), V(3)... mà chúng tôi sẽ nói sau.

10. Từ vựng phải được lập thức sao cho mỗi từ được phân định cho một trong các phạm trù từ vựng được phép đã liệt kê ở 7.

Ngữ pháp này dễ dàng được mở rộng. Chúng ta có thể mở rộng từ vựng:

<i>a</i>	Quán từ
<i>dog</i>	N
<i>under</i>	Giới từ
<i>lay</i>	V(1)

Chúng ta có thể bổ sung các quy tắc cú pháp. Chẳng hạn, *sat* (ngồi) và *lay* (nằm) đòi hỏi phải có một giới ngữ theo sau: *The cat lay under the table* (Con mèo nằm dưới bàn), nhưng không thể kết hợp trực tiếp với danh ngữ **the cat lay the mouse* (Con mèo nằm con chuột) hoặc một câu **the cat lay that the man chased the mouse* (Con mèo nằm người đàn ông bắt con chuột). Chúng ta có đặc trưng của V(1), tức là tiểu loại vị từ 1. Ngược lại, một vị từ như *caught* (bắt) đòi hỏi một danh ngữ theo sau: *The cat caught the mouse* (Con mèo bắt con chuột), nhưng không thể nói *the cat caught under table* (Con mèo bắt dưới bàn) hoặc *the cat caught that the mouse lay under the table* (Con mèo bắt con chuột nằm dưới bàn). Chúng ta sẽ đặc trưng hóa chúng là V(2). Vị từ *said* (nói) còn khác hơn nữa: nó đòi hỏi một câu theo sau: *The man said that the cat caught the mouse* (Người ta nói rằng con mèo bắt con chuột), nhưng không thể nói *the man said the cat* (Người ta nói con mèo) hoặc *the boy said under the table* (người con trai nói dưới bàn). Chúng ta gọi nó là thành viên của V(3). Để thích nghi với các tiểu loại vị từ khác nhau về ngữ pháp này, chúng ta có thể bổ sung thêm các quy tắc sau:

VP → V (2) NP

VP → V(3) S

Điều này sẽ kéo theo từ vựng bổ sung:

caught V(2)

chased V(2)

said V(3)

thought V(3)

Ngữ pháp được mở rộng chút ít có thể tạo sinh một số lớn các câu.

Thành tố cải biến bao gồm các quy tắc biểu diễn sự thay đổi của các chức năng.

Quy tắc cải biến là một quy tắc chỉ dẫn chuyển cấu trúc này sang cấu trúc khác. Quy tắc lấy một cấu trúc làm đầu vào và cho ra một cấu trúc khác, rõ ràng là sẽ cần có hai phần: *Phân tích cấu trúc* chỉ rõ đầu vào, cái cấu trúc mà quy tắc vận dụng vào và

sự thay đổi cấu trúc chỉ rõ cái là cấu trúc đầu ra.

Chúng ta sẽ quan tâm đến ba quy tắc: thứ nhất, quy tắc liên hệ các kiểu câu đặc biệt với nhau, như câu chủ động với câu bị động; thứ hai, một tập hợp các quy tắc miêu tả các thao tác hình thái học thuộc các loại khác nhau, như sự phù ứng về số giữa chủ ngữ và vị từ; cuối cùng, các quy tắc có trách nhiệm tạo sinh các câu phức.

Cái cấu trúc hiện ra sau hoạt động của tất cả các cải biến được hiểu là *cấu trúc cú pháp mặt* (syntactic surface structure). Về sau cấu trúc đó sẽ cần diễn ra các thành tố hình âm vị học (morphophonology) và âm vị học (phonology) để lĩnh hội hình thức âm vị học cuối cùng của nó.

Có thể tóm tắt cấu trúc toàn thể của mô hình này như nó được áp dụng vào một câu đơn

Quy tắc cấu trúc đoạn ngữ: “Cấu trúc cơ sở”

↓

Các cải biến

↓

Cấu trúc mặt

↓

(Các quy tắc âm vị học)

Trong mô hình này, tất cả các câu sẽ có ít nhất hai cấp độ miêu tả: *cấu trúc cơ sở* được tạo ra nhờ các quy tắc cấu trúc đoạn ngữ và *cấu trúc mặt* bắt nguồn từ hoạt động của các cải biến.

Có thể rút ra một số điều sau đây:

Một là, mô hình ngôn ngữ thứ nhất đã chú ý đặc biệt tới sự kiện là: ngôn ngữ là một tổ chức có tính cấu

trúc phức tạp. Tất cả các quy tắc mà chúng ta đã thấy khi hoạt động ở cấu trúc hay ở bộ phận của cấu trúc đều hoặc là phát triển hoặc là giảm bớt chúng. Đây là một trong các nét đặc trưng của ngôn ngữ loài người.

Hai là, mối quan hệ giữa cấu trúc cơ sở và cấu trúc mặt khiến chúng ta có thể lưu giữ nhiều sự khái quát hóa

được dẫn trong các đoạn mở. Như thế, quan hệ mô phỏng giữa các câu khác biệt trên bề mặt, chẳng hạn như một câu chủ động và dạng bị động tương ứng nảy sinh từ sự thực là cả hai phái sinh từ cùng một cấu trúc cơ sở. Ngược lại, một câu lưỡng nghĩa nảy sinh khi một sự phái sinh cải biến đã biến các cấu trúc cơ sở khác biệt thành một cấu trúc mặt duy nhất.

Cuối cùng, sự miêu tả này cho phép chúng ta nhận diện một lớp câu đặc biệt, *câu lõi (kernel sentence)*: các câu chủ động đơn giản, câu trần thuật, câu khẳng định. Nét khu biệt của các câu lõi thể hiện ở chỗ chúng là những câu được phái sinh với cơ cấu cải biến tối thiểu, chỉ với các cải biến bắt buộc. Như chúng ta đã thấy, các cải biến bắt buộc về bản chất là cải biến miêu tả phù ứng về số, trật tự bề mặt của dấu hiệu về thì và các thao tác giống như công việc nội bộ. Các câu khác - câu hỏi, câu phủ định, - sẽ chịu thêm một hoặc nhiều thao tác chuyển đổi cấu trúc không bắt buộc. Nhóm cải biến thứ ba là các cải biến chịu trách nhiệm tạo sinh các câu phức, các câu mà tự thân nó chứa đựng các câu hoặc các cấu trúc như câu với tư cách là các thành tố. Thí dụ: (1) *Kim said* (2) *that his mother expected him* (3) *to tell John* (4) *that...* (*Kim nói rằng mẹ nó muốn nó bảo John...*). Ở đây, các câu bao khác nhau được nhận diện là câu 1, câu 2... Rõ ràng, quá trình này rất sản sinh. Trong *Syntactic Structures*, hoạt động bao được biểu diễn bởi một tập hợp các cải biến khác biệt được gọi là các cải biến khái quát hóa, nó lấy hai cấu trúc câu làm đầu vào và sinh ra đầu ra là một cấu trúc duy nhất có

một câu được bao vào câu khác. Vấn đề này rõ ràng là một vấn đề quan trọng, nhưng cách giải quyết riêng biệt được chấp nhận trong *Syntactic Structures* là cực kì phức tạp, dẫn đến những khó khăn đáng kể về hình thức.

Có hai nhận xét về mô hình ngôn ngữ thứ nhất:

Thứ nhất, phải giải quyết mối quan hệ giữa cú pháp và ngữ nghĩa. Trong *Syntactic Structures*, Chomsky đã nhấn mạnh sự tự trị của cú pháp, đặc biệt với ngữ nghĩa. Vì thế, một câu như *Colorless green ideas sleep furiously* (*Những tư tưởng xanh lục không màu đang giận dữ ngủ*) cũng có thể được coi là hợp thức. Tuy nhiên, ông đã chú ý đến sự kiện rằng miêu tả một ngôn ngữ phải có các phương tiện để thảo luận mối quan hệ giữa cú pháp và ngữ nghĩa và chỉ ra rằng về phương diện này, các câu lõi phần nào có vai trò đặc quyền bởi vì nếu các câu lõi là các câu cơ sở, việc chúng được hiểu như thế nào là chìa khóa cho việc hiểu các câu nói chung.

Thứ hai, Chomsky coi ngôn ngữ như là một hệ thống hình thức của các quy tắc và điều đó dẫn ông tới phát hiện các đặc điểm toán học của các loại khác nhau của ngữ pháp hình thức. Một mình các quy tắc cấu trúc đoạn ngữ thì không đủ để miêu tả phạm vi các cấu trúc được tìm thấy trong một ngôn ngữ tự nhiên. Chẳng hạn, người ta nhận thấy một số cấu trúc được tìm thấy trong ngôn ngữ tự nhiên không thể được tạo sinh với các quy tắc cấu trúc đoạn ngữ.

Tóm lại, mô hình ngôn ngữ thứ nhất của ngôn ngữ học tạo sinh còn chưa hoàn chỉnh, còn nhiều vấn đề

chưa giải quyết được. Chomsky đã tiếp tục bổ sung, sửa chữa; từ đó chuyển sang pha thứ hai của ngôn ngữ học tạo sinh là: *Lí thuyết chuẩn hay Mô hình các bình diện*.

TÀI LIỆU THAM KHẢO

1. Noam Chomsky, *Những chân trời mới trong nghiên cứu ngôn ngữ và ý thức*, (Hoàng Văn Vân dịch), Nxb GD, H., 2007.

2. Nguyễn Đức Dân,

a. *Ngữ pháp tạo sinh*, Trong *Ngôn ngữ học: khuynh hướng, khái niệm, lĩnh vực*, Nxb KHXH, H., 1984, tr.96 - 119.

b. *Chomsky Noam*, Trong *Ngôn ngữ học: khuynh hướng, khái niệm, lĩnh vực*, Nxb KHXH, H., 1984, tr.168 - 172.

c. *Avram Noam Chomsky: “người có trí tuệ nhất thế giới”*, T/c Từ điển học và Bách khoa thư, Số 5, 2011.

3. Nguyễn Thiện Giáp,

a. *Giáo trình ngôn ngữ học*, Nxb ĐHQG, H., 2008.

b. *777 khái niệm ngôn ngữ học*, Nxb ĐHQG, H., 2010.

c. *Ngôn ngữ học tạo sinh của N.Chomsky: Đối tượng và mục đích*, T/c Ngôn ngữ, Số 1, 2012.

4. Lưu Nhuận Thanh, *Các trường phái ngôn ngữ học phương Tây*, (Đào Hà Ninh dịch), Nxb Lao động, H., 2004.

5. *Những cơ sở triết học trong ngôn ngữ học* (Trúc Thanh dịch), Nxb GD, H., 1984.

6. R.H. Robins, *Lược sử ngôn ngữ học*, (Hoàng Văn Vân dịch), Nxb ĐHQG, H., 2003.

7. *The Linguistics, Encyclopedia*, Edited by Kirsten Malmkj Ær, London and New York, 1995.