

La linguistique fonctionnelle et structurale de Prague
et sa continuation à l'époque
de la description formelle

Petr SGALL

Université Charles, Prague

LA MÉTHODE DE BASE de la linguistique est l'analyse empirique, qui ne peut que partir des données du discours, de la communication verbale. Seuls les phénomènes de l'expression du discours sont directement observables et leur analyse est l'unique moyen utilisable pour discerner des traits du système de la langue. Mais cela ne signifie pas que la notion de langue serait un cas d'hypostase. De ce point de vue, la langue peut être regardée comme un objet abstrait, construit (cf. Lieb, 1983). D'un autre côté, la langue comme système appartenant à la réalité psychologique peut être modelée par des procédés tirés du connectionnisme (cf. Schnelle, 1991).

Pour l'école linguistique de Prague le caractère abstrait de la langue était un fait de base, un des points de départ principaux de sa méthode structurale depuis le début des activités de l'école. La langue constitue, du point de vue saussurien, un système d'oppositions dans lequel la substance des éléments individuels n'est pas d'une signification primaire.

Simultanément, le fonctionnalisme pragois considérait la langue comme un véhicule de la communication et rendait compte des rapports entre le signifiant et le signifié dans le système des signes qu'il voyait dans la langue. C'est ici que l'on peut voir la source de la méthode de comparaison, qui avait tant d'importance pour l'école de Prague dans sa période classique. L'un des points décisifs pour établir une méthode générale, c'était la possibilité et la nécessité de comparer des langues entre elles. Pour Vilém Mathesius (1936), la comparaison des langues constituait une voie pour identifier les traits de la structure sémantique (profonde) des langues, dans laquelle les langues différentes ne sont pas tellement éloignées l'une de l'autre, comme aux niveaux des phonèmes et des morphèmes. La comparaison du tchèque à l'anglais

était pour lui, et aussi pour Jan Firbas, qui continue ses efforts, une base appropriée pour discerner les divers moyens et fonctions de l'articulation de la phrase en thème et rhème.

De même, la formulation des lois générales par Roman Jakobson (1958) sous forme d'implication est basée sur la comparaison des diverses langues et offre une nouvelle perspective à cette comparaison. La signification de ces lois pour la méthodologie linguistique a été soulignée par Holenstein (1985 : 207).

A l'intérieur d'une seule langue, il faut comparer les différents choix de signifiés en présence d'un seul signifiant (même avec les « invariants » de Jakobson (1936), on doit aussi chercher et décrire les significations contextuelles et les cas d'homonymie). Il en est de même de l'autre direction de l'analyse, menant de la signification à l'expression; c'était cette orientation de l'analyse, le point de vue du locuteur, qui a tout d'abord été soulignée par Mathesius.

La formulation nouvelle et sobre de la typologie classique dans le contexte de la linguistique structurale par Vladimír Skalička (1935, 1979) est un autre exemple de l'importance attribuée à la comparaison des langues et aussi à l'examen des choix possibles de formes d'expression. Je crois qu'en employant cette typologie, on peut identifier les qualités fondamentales des types individuels de langues. Ces qualités concernent les limites étroites du choix des formes de l'expression des valeurs grammaticales. Tandis que les valeurs lexicales sont toujours exprimées par des morphes (des chaînes de phonèmes), pour les unités grammaticales, le langage, avec sa base acoustique et articulatoire, ne donne que le choix suivant :

- (i) morphes pareils aux unités lexicales : mots accessoires, type analytique (le français, l'anglais);
- (ii) morphes associés aux bases lexicales : affixes, agglutination (langues turco-tatares, finno-ougriennes etc.);
- (iii) altérations de la fin de la base : une seule désinence à chaque forme de mot, la flexion (langues indo-européennes classiques, langues slaves);
- (iv) altérations au centre de la base : introflexion (langues sémitiques);
- (v) ordre des morphes lexicaux uniquement : polysynthèse ou isolation (le chinois).

Cette base permet d'établir une typologie plus riche que celle basée sur l'ordre des mots et des affixes (cf. Sgall, 1986). Ces fondements méthodologiques reposent sur les lois d'implication de Jakobson, lesquelles ont obtenu, chez Skalička, un caractère probabiliste.

Un autre aspect important du programme de linguistique synchronique, né ici au bord du lac Léman et qui n'est pas encore, de loin, réalisé, est la distinction entre la forme du contenu et sa substance (dans la terminologie de Ferdinand de Saussure). A Prague, non seulement les fondateurs de l'Ecole, mais aussi Pavel Trost, Miloš Dokulil et d'autres penseurs appartenant à sa deuxième génération ont décidé de travailler en linguistique avec des critères opérationnels pour pouvoir distinguer ce qui appartient vraiment à la structure (forme) de la langue, et ce qui n'est qu'un trait du contenu ontologique ou cognitif.

Dans le groupe de recherches linguistiques théoriques de l'Université Charles, nous nous sommes occupés de rechercher de tels critères pour pouvoir établir les unités présentes dans la structure syntaxique (profonde) de la phrase, ou, en d'autres termes, au niveau de la signification, de la forme linguistique du contenu (cf. Sgall, Hajičová et Panevová, 1986).

Quant au noyau de cette structure, l'ensemble des types de modifications (des positions de valence dans la terminologie de Lucien Tesnière), on peut illustrer ce procédé par les exemples suivants :

- (a) D'après l'un des critères connus, le cas de l'homonymie d'une forme d'expression (présence de deux signifiés) peut être distingué de la présence d'un seul signifié vague : le locuteur est toujours capable de résoudre l'homonymie de sa propre formulation, ce qui veut dire que s'il ne peut pas préciser le contenu de ses paroles on est en présence d'un élément vague. Employant ce critère, on peut dire que la distinction entre des unités de la substance de contenu telles que l'Agentif et l'Expérienceur ne peut qu'appartenir au domaine cognitif; ce qui est présent au niveaux de la signification linguistique (la forme du contenu), c'est l'unité vague (mais non homonyme) de l'Acteur, du premier actant de Tesnière (1959). S'il est possible qu'en prononçant (1) le locuteur ne sache pas s'il s'agissait d'une action délibérée, cela veut dire qu'une signification vague est présente, et non une expression homonymique.

(1) Marcel a cassé la fenêtre.

Dans d'autres cas, le locuteur sait la différence, mais cela est conditionné par sa déduction; il peut tirer les conséquences de la connaissance qu'il a de certains faits, ou du contexte; avec d'autres verbes il peut s'agir du contexte intraphrastique, de la signification du verbe lui-même.

- (b) Un autre critère de cette nature (établi par J. Panevová, cf. Sgall, Panevová et Hajičová, 1986 : chap. 2) peut s'employer pour distinguer entre une position de valence vacante et un cas d'ellipse (dans lequel un élément présent dans la signification de la phrase est supprimé dans sa forme de surface). Le locuteur est toujours capable de résoudre une ellipse. Cela veut dire que si A a prononcé (2) et son interlocuteur lui pose la question (3), il est impossible que A réponde par (4).

(2) Marie est arrivée. (3) Où est-elle arrivée? (4) Je ne sais pas.

- (c) Dans la syntaxe, une des questions importantes est celle de l'orientation de la relation de dépendance. Pour décider lequel des deux éléments d'un syntagme représente le mot principal et lequel est le mot dépendant, on peut se servir du critère de la suppression syntaxique, si l'on ne reste pas au niveau des mots individuels, mais passe par celui des classes de mots (parties du discours). Par exemple, dans une paire substantif-adjectif, l'adjectif peut manquer, donc c'est le substantif qui régit le syntagme. C'est pareil avec une paire verbe-objet: il y a des verbes qui exigent la présence d'un objet obligatoire, mais au niveau plus abstrait, c'est le verbe qui est toujours nécessaire dans la structure de la phrase, et non l'objet.
- (d) Il est aussi possible de trouver un critère pour identifier la présence de la synonymie (au sens d'une seule unité du niveau de la signification linguistique correspondant à deux formes d'expression différentes). Deux signifiants peuvent être considérés comme vraiment synonymes (au sens linguistique, plus étroit qu'une équivalence sémantique au sens de la logique) seulement si l'on peut remplacer l'une d'eux par l'autre dans un contexte quelconque, sauf les contextes de citation. Bien que (5) et (6) soient équivalents au sens logique (correspondants à la même proposition Carnapienne, étant vrai par rapport à la même classe des mondes possibles), d'après ce critère ils ne sont pas synonymes, car il existe des

contextes tels que dans (7) et (8), où les deux propositions correspondantes ne sont pas identiques.

(5) Le livre de Paul a été publié par Hachette.

(6) Hachette a publié le livre de Paul.

(7) Des livres sur ce sujet sont été publié tout d'abord par Hachette.

(8) Hachette a publié tout d'abord des livres sur ce sujet.

Un tel critère nous permet aussi de caractériser (9) et (10) comme n'étant pas synonymes, bien qu'il correspondent à la même proposition (c'est-à-dire qu'ils sont valables pour tous les mondes possibles).

(9) Deux et deux font quatre. (10) Trois et trois font six.

Il est important que la signification linguistique, comme aussi la représentation d'une proposition dans des systèmes de la logique d'intension (tant celui de Tichy, 1990), constitue une classification des sens plus détaillée que l'intension elle-même. On peut voir que ces critères, importants pour l'identification des éléments syntaxiques et sémantiques de la structure linguistique, sont aussi d'une grande importance pour une coopération de la linguistique avec la sémantique d'intension, pour les problèmes des attitudes « propositionnelles » etc. La structure profonde (c'est à dire le niveau de la signification linguistique) peut être comprise comme un niveau dépourvu de l'ambiguïté au sens de Richard Montague ou de David Lewis. Avec ce niveau, on obtient une base pour une division du travail entre la linguistique et le domaine interdisciplinaire de l'interprétation sémantique, où il est nécessaire de combiner les efforts de la linguistique avec ceux de la logique et des sciences cognitives. C'est précisément une telle coopération qui peut aboutir à une description intégrale de la sémantique du langage humain (y compris sa structuration du point de vue de la pragmatique, avec les expressions indexiques, avec les catégories de temps, de mode, d'aspect verbal, etc.).

Un emploi rigoureux de tels critères nous permet de présenter une approche de la linguistique théorique qui a les deux qualités importantes suivantes :

- (i) elle correspond aux exigences rigoureuses de la méthodologie de la présente linguistique formelle, dont la base syntaxique a été créée par N. Chomsky et dont le traitement de la sémantique a été élaboré par R. Montague, B. H. Partee et autres;
- (ii) elle emploie des idées du structuralisme européen classique, lesquelles, au moins quant au domaine de la syntaxe, sont mal connues de la linguistique américaine (à de rares exceptions près, telles que la grammaire des cas de Fillmore).

Le système de description de la langue présenté par Sgall, Hajičová et Panevová (1986) contient des moyens pour une spécification formelle de la structure syntaxique (profonde) de la phrase, des procédés venant de cette structure jusqu'à la forme phonétique des phrases, et aussi pour une caractérisation de l'interprétation sémantique au sens de l'analyse logique du langage (qui a un caractère interdisciplinaire). On peut voir que les représentations de la structure des phrases obtenues par cette méthode, qui n'ignore pas les exigences de F. de Saussure et de L. Tesnière, sont beaucoup plus simples et naturelles que les représentations basées sur les « constituants immédiats » de Chomsky. Un autre avantage est la possibilité de décrire, aussi dans ce cadre, l'articulation de la phrase en thème et rhème d'une manière systématique et transparente. Cela nous permet de traiter la phrase non seulement comme une structure abstraite, mais comme une forme de l'unité de base appartenant à un contexte, servant dans le procès de communication.

L'orientation de la théorie et de la méthode linguistiques vers le rôle de la langue dans la communication est un postulat qui est tout à fait actuel et urgent aujourd'hui. Il est clair qu'on doit décrire la langue comme un système d'interaction, qui est le moyen principal de la communication humaine et dont la structure a subi de l'influence du fait que son développement millénaire a eu lieu dans les conditions de la communication. Le point de vue téléologique de l'école de Prague ne peut que souligner que les qualités caractéristiques de la langue doivent leur origine aux exigences des participants du discours et aux conditions de la communication.

Pour pouvoir décrire non seulement la sémantique (sémantico-pragmatique) de la phrase, mais aussi celle du discours, il faut (i) avoir une description de la phrase prise dans son contexte, et aussi (ii) aboutir à un modèle de la représentation d'une chaîne des énoncés (c'est à dire des occurrences de phrases).

- (i) La première de ces exigences peut être assurée sur la base d'une description de l'articulation en thème et rhème (information « donnée » et « nouvelle »), si cet articulation est vue comme une des structurations dans la représentation profonde de la phrase. Dans le livre cité plus haut nous avons présenté une approche qui peut remplir cette tâche pour différentes langues.
- (ii) Le second point concerne de questions telles que celles de la cohésion du discours. Un des aspects importants pour le traitement de la sémantique d'une chaîne d'énoncés, c'est l'ensemble des relations anaphoriques et co-référentielles. Il est nécessaire d'identifier un mécanisme fini permettant à l'interlocuteur de spécifier la référence des expressions employées dans un discours, même quand les possibilités de leur référence à divers objets du contenu ne sont pas restreintes. Une occurrence du mot « table » peut se référer à une table quelconque (réelle ou irréelle, spécifiée ou non) ou à toute la classe des tables (d'une manière générique), mais l'interlocuteur peut, en principe, employer des moyens restreints pour déterminer la table dont on lui parle (dans un discours bien formulé). Le mécanisme dont il se sert est basé sur la notion des degrés de saillance (actualisation) des entités mentionnées dans le discours (des référents), cf. Hajičová et Vrbová (1981). Dans la règle, une entité est mentionnée tout d'abord dans le rhème d'un énoncé, ce qui lui donne un degré de saillance maximal. Puis, la même entité peut être mentionnée une autre fois, disons dans l'énoncé suivant, probablement dans son thème (comme une pièce de l'information « donnée » ou « connue »), ce qui soutient sa saillance. Si, dans la chaîne d'énoncés, on en trouve un dans lequel cet entité n'est pas mentionnée, cela diminue sa saillance et rend probable le fait qu'elle devra être exprimée par un moyen plus fort qu'un pronom faible si elle apparaît par la suite dans le discours (par le nom, ou même par tout un groupe nominal, s'il faut distinguer p. ex. une table noire d'une autre table dont on a parlé aussi dans la même chaîne d'énoncés).

Un tel modèle devrait être combiné avec les moyens de description employés dans les approches de I. Heim (« file change semantics »), de H. Kamp (« discourse representation theory ») et d'autres auteurs, appartenant tout d'abord à l'école de Barbara Partee, si on veut aboutir à une description au niveau de la linguistique formelle. Un tel procédé

semble être le plus approprié pour rendre compte du contenu de l'énoncé non seulement comme une pièce d'information propre au « locuteur idéal » de Chomsky, mais comme une opération changeant l'état psychique de l'interlocuteur. Cela constitue le point principal du programme actuel de la linguistique théorique, qui intègre la description de la langue dans un modèle interactif du processus communicationnel.

Si je peux ajouter une remarque concernant l'emploi d'ordinateurs dans la théorie linguistique (non seulement dans des applications), je voudrais souligner que d'une part un modèle complexe de la grammaire, de la structure du discours ou de sa sémantique ne peut être contrôlé qu'en employant un système de génération de phrases (et de leur chaînes) aléatoire. D'autre part, une approche complexe basée sur des méthodes dont nous avons parlé a permis, en principe, de construire un modèle automatique correspondant à certains traits du procédé de compréhension d'un texte (cf. Hajičová et Sgall, 1984). Il est possible de modeler non seulement le procédé de l'interlocuteur à partir de la forme phonétique ou graphique d'un texte jusqu'à sa structure profonde, mais aussi un ensemble de régularités dans le domaine du cognitif. Dans des expérimentations avec un système automatique donnant des réponses aux questions sur la base d'un texte technique traité en entrée, on a pu élaborer quelques dizaines de types de règles d'inférence, ce qui peut être illustré par les exemples suivants (même si les expérimentations pragoises concernent des phrases tchèques) :

- (11) On peut exécuter par ces moyens les opérations arithmétiques et algébriques.
- (12) On peut exécuter par ces moyens les opérations algébriques.
- (13) Parce qu'on a ces moyens, on peut travailler effectivement.
- (14) On a ces moyens.
- (15) Par ces moyens on peut décrire des phrases anglaises simples.
- (16) Par ces moyens on peut décrire des phrases anglaises.
- (17) Les phrases anglaises simples peuvent être décrites par ces moyens.
- (18) Les phrases anglaises peuvent être décrites par ces moyens.
- (19) On appelle conjugaison la partie de la morphologie concernant le verbe.
- (20) En français, la partie de la morphologie concernant le verbe appartient aux aspects les plus complexes de la grammaire.

(21) Est-ce qu'en français la conjugaison est complexe ?

Les règles d'inférences formulées dans notre système d'expérimentations permettent de déduire (12) de (11), où il s'agit d'une distribution de la relation de la conjonction; des cas comme (14) et (13) une proposition subordonnée exprimant la cause peut être (dans certaines conditions) rendue indépendante de la proposition principale. Il est aussi possible de déduire (16) de (15), ce qui peut être important pour pouvoir répondre à des questions comme (22); il est possible de réduire ici le rhème de la phrase (16), contenant un élément indéfini. D'autre part, le thème (qui contient régulièrement des éléments définis) ne peut être réduit, ce qui veut dire que (18) ne peut être construit sur la base de (17) par notre système. Le traitement de définitions est illustré par (19), dont la présence dans le texte permet de répondre par (20) (si une telle phrase est présente dans le texte) à la question (21).

(22) Est-ce que des phrases anglaises peuvent être décrites avec ces moyens ?

Le sens de ces remarques est de montrer une possibilité d'employer des résultats de la linguistique structurale européenne dans le contexte de la méthodologie de la théorie linguistique moderne et dans la perspective d'une coopération de la linguistique avec la logique et la science cognitive.

© Petr Sgall

RÉFÉRENCES BIBLIOGRAPHIQUES

- HAIJČOVÁ, E., SGALL, P. (1984) *Contributions to functional syntax, semantics and language comprehension*. Amsterdam : Benjamins; Prague : Academia.
- HAIJČOVÁ, E., VRBOVÁ, J. (1981) « On the salience of the elements of the stock of shared knowledge ». In *Folia linguistica*, 15, p. 291-303.
- HOLENSTEIN, E. (1985) *Sprachliche Universalien. Eine Untersuchung zur Natur des menschlichen Geistes*. Bochum : Brockmeyer.
- JAKOBSON, R. (1936) « Beitrag zur allgemeinen Kasuslehre ». In *Travaux du Cercle linguistique de Prague*, 6, p. 240-288.
- (1958) « Typological studies and their contribution to historical comparative linguistics ». In *Proceedings of the Eighth International Congress of Linguists*. Oslo, p. 17-35.
- LIEB, H.-H. (1983). *Integrational linguistics*. Amsterdam : Benjamins.
- MATHESIUS, V. (1936) « On some problems of the systematic analysis of grammar ». In *Travaux du Cercle linguistique de Prague*, 6, p. 95-107.
- SCHNELLE, H. (1991) *Natur der Sprache. Die Dynamik der Prozesse des Sprechens und Verstehens*. Berlin : W. de Gruyter.
- SGALL, P. (1986) « Classical typology and modern linguistics ». In *Folia linguistica*, 20, p. 15-28.
- HAIJČOVÁ, E., PANEVOVÁ, J. (1986) *The meaning of the sentence in its semantic and pragmatic aspects*. In J. L. MEY (éd.). Dordrecht : Reidel; Prague : Academia.
- SKALIČKA, V. ([1935] 1979) *Zur ungarischen Grammatik*. Prague, p. 59-125.
- (1979) *Typologische Studien*. In HARTMANN, P. (éd.). Braunschweig-Wiesbaden : Vieweg.
- TESNIÈRE, L. (1959) *Eléments de syntaxe structurale*. Paris : Klincksieck.
- TICHY, P. (1990) *Foundations of Frege's logic*. Berlin : W. de Gruyter.