

Manuel

**de linguistique française
et de linguistique générale**

Samuel Bidaud



Univerzita Palackého
v Olomouci

Univerzita Palackého v Olomouci

Manuel

de linguistique française
et de linguistique générale

Samuel Bidaud

Olomouc 2020

Oponenti/Relecteurs :

Mgr. Ľudmila Lacková, Ph.D.

Mgr. Jan Zatloukal, Ph.D.

Zpracování a vydání publikace bylo umožněno díky finanční podpoře, udělené roku 2020 Ministerstvem školství, mládeže a tělovýchovy ČR v rámci Fondu rozvoje UP, Filozofické fakultě Univerzity Palackého v Olomouci.

La préparation et la publication de cet ouvrage ont été rendues possibles grâce à un soutien financier accordé en 2020 par le Ministère de l'Éducation, de la Jeunesse et du Sport de la République tchèque à la Faculté des Lettres de l'Université Palacký d'Olomouc, dans le cadre du Fond pour le développement UP (FRUP).

Neoprávněné užití tohoto díla je porušením autorských práv a může zakládat občanskoprávní, správněprávní, popř. trestněprávní odpovědnost.

L'utilisation non autorisée de cet ouvrage constitue une violation des droits d'auteur et peut engendrer des poursuites sur le plan civil, juridique et éventuellement pénal.

1. vydání/1^{ère} édition

© Samuel Bidaud, 2020

© Univerzita Palackého v Olomouci, 2020

ISBN 978-80-244-5799-4 (print)

ISBN 978-80-244-5800-7 (online: PDF)

Table des matières

Préface	7
Chapitre 1	
Qu'est-ce qu'une langue ?	9
Chapitre 2	
Quelques notions fondamentales de linguistique	
Le signe – La langue et la parole – La synchronie et la diachronie	15
La langue et la parole	15
Le signe linguistique	17
La synchronie et la diachronie	20
Chapitre 3	
La communication	23
Les participants de la communication	23
Les fonctions du langage	24
Chapitre 4	
Phonétique et phonologie	31
Phonétique et phonologie	31
L'Alphabet Phonétique International	32
Comment classer les phonèmes ?	36
Chapitre 5	
La morphologie	43
Les morphèmes	43
La formation des mots	46

Chapitre 6

La morphosyntaxe	57
Les parties du discours	57
Les fonctions	72
La syntaxe de Tesnière	75
Pour conclure	87

Chapitre 7

La sémantique	95
Les relations sémantiques à l'intérieur du lexique	95
La sémantique structurale	97
La sémantique textuelle	98

Chapitre 8

La variation du français	107
Les registres de langue	107
Le français oral, le français écrit et l'« écrit transmis »	111
La variation dans l'espace	112

Chapitre 9

Comment classer les langues ?	121
La classification historique des langues	121
La classification typologique	125

Conclusion	129
-------------------------	------------

Bibliographie	133
----------------------------	------------

Préface

Ce manuel s'adresse prioritairement aux étudiants et étudiantes de philologie française. Il répond à un double objectif : d'une part, il souhaiterait présenter l'ensemble des grands domaines de la linguistique française, et d'autre part donner un panorama des notions les plus importantes de linguistique générale. Il est en effet impossible d'étudier la première sans un minimum de connaissances concernant la seconde. Cet ouvrage se propose ainsi d'aborder non seulement des champs comme la phonétique, la morphologie ou la syntaxe du français, mais il a également pour but d'introduire à des problématiques plus larges, qu'il s'agisse du concept de langue ou de signe linguistique, ou encore des fonctions du langage, pour citer trois exemples parmi d'autres. Enfin, il se veut également pratique et illustratif. En ce sens, la majorité des chapitres comprennent un certain nombre d'exercices qui permettront aux lecteurs et lectrices de vérifier leur compréhension des différents points étudiés. Nous avons en outre fréquemment ajouté, à un ensemble d'exercices corrigés, plusieurs exercices non corrigés pour aller plus loin.

La linguistique a parfois la réputation d'être une discipline un peu sèche. Nous aimerions faire apparaître qu'il n'en est rien dès lors qu'on ne l'envisage pas comme une discipline purement technique mais comme un moyen d'accéder à l'homme présent derrière la langue, comme le disait le stylisticien Leo Spitzer. La linguistique est en effet un voyage à travers les mots et les visions du monde que l'on retrouve derrière eux, de sorte que chaque langue porte en elle tout un univers. Karel Čapek, dans son *Éloge de la langue tchèque*, soulignait que l'écrivain a la capacité de redécouvrir les mots dans leur dimension originelle ; et c'est cette capacité à voir la langue comme autre chose qu'un simple instrument de communication pour les besoins de la vie courante qui caractérise la poésie. Ainsi l'étude de la langue rencontre-t-elle finalement celle de la littérature.

Nous espérons que ce manuel pourra non seulement aider les étudiants et étudiantes des cursus de philologie française, mais aussi et surtout qu'ils et elles trouveront du plaisir à étudier la langue française et, au-delà, la linguistique dans son ensemble.

Nous tenons à remercier Aneta Křemenová, Ľudmila Lacková, Michaela Vodová, Stanislav Pisklák et Jan Zatloukal, qui ont bien voulu relire un ou plusieurs chapitres de ce livre. Enfin, nous remercions l'Université Palacký d'Olomouc, qui a permis la rédaction et la publication de ce manuel à travers le projet FRUP_2020_007.

Chapitre 1

Qu'est-ce qu'une langue ?

Intuitivement, la question de savoir ce qu'est une langue ne se pose guère. On « sait » par exemple que le français, le tchèque, l'italien ou le suédois sont des langues. Pourtant, lorsque l'on se penche de plus près sur le sujet dans le but de proposer un ensemble de critères qui permettent de définir la notion de langue, de nombreuses difficultés surgissent.

Si l'on consulte un dictionnaire de référence de la langue française, le *Trésor de la Langue Française informatisé* [TLFi], on peut y lire la définition suivante de la langue dans la rubrique [La langue comme système] :

Système de signes vocaux et/ou graphiques, conventionnels, utilisé par un groupe d'individus pour l'expression du mental et la communication.

Analysons de plus près cette définition point par point afin de voir ce qu'elle implique.

a) La langue est un système de signes

L'idée de système suppose que les signes que nous utilisons sont régis par un principe d'organisation à différents niveaux : phonétique, morphologique et syntaxique. Prenons l'exemple d'un mot comme *fleur* :

- ▶ phonétiquement, il correspond à la suite de sons *f-l-eu-r*. Ces derniers font l'objet d'un choix à l'intérieur d'un nombre restreint d'unités, qui forment l'ensemble des sons d'une langue donnée, ici le français. Ces sons, en tant qu'unités abstraites, sont appelés *phonèmes* (voir *Chapitre 4. Phonétique et phonologie*). Ils sont en nombre fini et s'opposent entre eux à partir de leurs traits articulatoires : par exemple, /f/ et /v/ ont le même lieu et le même mode d'articulation, mais l'un, /v/, est sonore (les cordes vocales vibrent quand on le prononce), alors que l'autre, /f/, est sourd (les cordes vocales ne vibrent pas). Autrement dit, /f/ et /v/ s'opposent à partir de l'absence ou de la présence de sonorité ;
- ▶ au niveau de la morphologie, on peut dire que le nom *fleur*, en français, est féminin et qu'il est au singulier. Là aussi, ces deux catégories reposent sur un système et sur un choix à l'intérieur de ce dernier : ainsi, en français, le système du genre oppose le féminin au masculin, et celui du nombre le singulier au pluriel ;

- ▶ enfin, le nom *fleur* entraîne un phénomène d'accord lorsqu'il est intégré dans une phrase : dans *Cette fleur est très belle*, le déterminant *cette* et l'adjectif *belle* sont au féminin puisque *fleur* est un nom féminin, et le verbe *être* est au singulier puisque *fleur* est lui aussi au singulier.

Comme il ressort de ce qui précède, la langue est donc un système de signes régi par des règles de fonctionnement précises. Le locuteur, à chaque fois, effectue un choix à l'intérieur d'un *paradigme*, c'est-à-dire d'un ensemble de possibilités qui se présentent à lui : le nom est ou masculin ou féminin¹, ou singulier ou pluriel ; le verbe est ou à l'indicatif, ou au subjonctif, ou à l'infinitif, etc. ; il est ou au présent, ou au futur, ou à l'imparfait, etc. Les choix qui sont faits sur l'axe paradigmatique entraînent des conséquences au niveau de la combinaison des signes entre eux dans la phrase, par exemple à travers le phénomène de l'accord : si je choisis un nom au pluriel, en effet, le verbe qui s'accorde avec ce dernier sera normalement au pluriel. On parle alors, pour ce second niveau d'analyse, qui est horizontal, d'axe *syntagmatique*.

Les éléments qui relèvent de la phonétique, de la morphologie et de la syntaxe sont en nombre fini. On ne peut pas ajouter un troisième genre en français au masculin et au féminin, par exemple, d'où le fait que le genre est un système *fermé*. En revanche, le lexique est un système ouvert : des mots comme *ciel*, *mer*, *rocher*, *arbre* ou *soleil* sont en principe illimités. Ils obéissent toutefois eux aussi à un principe de classement, mais ce dernier est beaucoup plus vague : les mots que nous avons cités, par exemple, appartiennent à la classe des /éléments naturels/.

De façon plus générale, on peut dire que la langue repose sur un principe de **double articulation**. Ce principe, qui a été théorisé par le linguiste français André Martinet (1980), signifie que toute langue est composée de deux types d'unités : les unités de première articulation, appelées monèmes ou unités minimales de sens, pour lesquelles nous préférons ici le terme *morphèmes* (voir *Chapitre 5. La morphologie*), et les unités de seconde articulation, les *phonèmes*. La combinaison des premiers et des seconds répond à un principe d'économie du langage.

Prenons la phrase *Mon amie arrive demain*. On a dans cette phrase un certain nombre de morphèmes : *mon* contient le morphème du masculin, celui du singulier et celui de l'idée de possession ; *amie* contient le morphème du féminin, marqué par le graphème *e* à l'écrit, celui du singulier et celui du sens de *ami* (« être qui est cher ») ; *arrive* contient le morphème de l'indicatif, celui du présent (qui prend ici une valeur de futur proche du fait de l'adverbe *demain*), celui de la troisième personne et celui du singulier, ainsi que l'idée d'arriver (« fin

¹ En réalité il n'est permis au locuteur de choisir le genre d'un mot que dans les cas de motivation « réelle » : *un étudiant vs. une étudiante*, *un professeur vs. une professeure*, etc. Lorsque le genre n'est pas motivé extérieurement et qu'il s'agit d'une donnée de départ associée au mot, comme dans *fleur* qui est forcément de genre féminin ou dans *bateau* qui est forcément de genre masculin, le locuteur ne « choisit » pas entre plusieurs possibilités.

de déplacement en direction d'un point ») ; *demain* enfin contient le morphème de l'idée de futur proche (demain = « après le jour d'aujourd'hui »).

À côté de ces morphèmes, on distingue également des phonèmes, qui sont les suivants : *m-on-n-a-m-i-a-r-i-v-d-e-m-ain*. On a donc 14 phonèmes, que l'on peut transcrire de la façon suivante en Alphabet Phonétique International : /mɔ̃namiaʁivdɛmɛ̃/.

Le principe d'une langue est de pouvoir produire, à partir d'un nombre fini de phonèmes, un nombre théoriquement illimité de signes. Ainsi, le français, à partir d'un nombre restreint de phonèmes (entre 34 et 36 selon les variétés régionales), permet de créer un nombre infini de signes (le TLFi, à titre d'exemple, répertorie 100 000 mots, mais il n'est pas exhaustif et l'on peut par ailleurs toujours créer des mots nouveaux, les *néologismes*).

b) Les signes sont vocaux et/ou graphiques

On pourrait penser que toute langue dispose d'une écriture, or c'est loin d'être toujours le cas. C'est pour cette raison que l'on peut établir une distinction entre les langues qui ont un système de signes vocaux et graphiques, et celles qui n'ont qu'un système de signes vocaux et qui restent des langues purement orales. Ainsi le français, le tchèque, etc., sont des langues qui sont à la fois vocales (orales) et écrites. Parmi les langues qui sont écrites, beaucoup ne le sont que depuis une époque récente : c'est le cas, en Europe, de l'estonien ou du roumain, qui ne sont écrits que depuis le 16^{ème} siècle. Au contraire, l'ulau-sain, parlé en Papouasie Nouvelle-Guinée, l'aasáx, parlé en Tanzanie, ou le mixtec-chazumba, parlé au Mexique, ne sont pas écrits et sont des langues purement orales. Certaines langues éteintes, enfin, comme l'étrusque, ne nous sont connues que par les documents écrits qu'elles nous ont laissés, mais il s'agissait toujours de langues qui étaient aussi et avant tout des langues orales. En ce sens, il n'existe pas de langue qui soit uniquement écrite.

c) Les langues utilisent un système de signes conventionnel

Le fait que les signes qui composent les langues sont conventionnels implique que la langue a une dimension collective et qu'aucun locuteur, en soi, ne peut la changer par sa propre volonté, au risque de n'être plus compris de la communauté linguistique à laquelle il appartient. Le système de signes est reçu par les locuteurs comme tel : il s'agit donc pleinement d'une *convention*.

d) Les langues sont utilisées par un groupe d'individus

La langue présuppose une situation d'interlocution, c'est-à-dire d'avoir un locuteur et un interlocuteur. Elle est en effet un moyen de communiquer et naît dans le cadre de l'interaction sociale entre plusieurs individus.

Il y a toutefois des langues qui sont mortes, et donc qui ne sont plus utilisées, et des langues qui sont parlées par un groupe très restreint d'individus, voire par un seul individu. Ce fut

le cas de l'oubykh, une langue caucasienne, dont le dernier locuteur, Tevfik Esenç, est mort en 1992, ou, un siècle plus tôt, du végliote, une variante du dalmate (langue romane) qui n'était plus parlé que sur l'île de Krk (proche de la Croatie) et dont le dernier locuteur, Tuone Udaina, est décédé en 1898. Mais il ne faut pas oublier que ces langues ont été à l'origine parlées par des communautés plus larges, et qu'elles remplissaient donc le critère d'être partagées par un ensemble d'individus entre lesquels elles permettaient la communication.

e) La langue sert pour l'expression du mental et la communication

Il s'agit là du but de la langue, qui est faite pour transmettre une expérience au sens large, d'un locuteur A vers un interlocuteur B, lequel devient à son tour locuteur dans le cadre de l'échange. La langue peut, nous le verrons, avoir des fonctions multiples, mais dans la majorité des cas elle permet le *dialogue*, quand bien même il s'agirait du langage intérieur, où le locuteur s'adresse en réalité à lui-même.

Au terme de cette brève analyse, nous avons réussi à dégager un ensemble de critères qui permettent de définir l'entité « langue » d'un point de vue interne. Mais, malgré tout, ces critères sont-ils exhaustifs ? Nous permettent-ils, par exemple, de pouvoir dire que le français est une langue mais que le français parlé en Bourgogne n'en est pas une, que le tchèque est une langue mais que le tchèque parlé en Silésie n'en est pas une, ou bien que le tchèque et le slovaque sont deux langues différentes ? Il nous faut, à cette fin, faire intervenir des critères non plus internes mais externes.

Nous partirons ici du critère de l'intercompréhension. On pourrait penser que, dès lors que deux locuteurs d'une langue A et d'une langue B ne se comprennent pas, alors nous sommes en face de deux langues différentes, et que, inversement, lorsque ces mêmes locuteurs se comprennent complètement, ils parlent dès lors la même langue. Ainsi un locuteur du chinois et un locuteur du français ne se comprennent pas si chacun d'eux parle sa propre langue : le chinois et le français sont donc bien deux langues différentes. De même un locuteur du portugais avec un locuteur de l'hindi, un locuteur du danois avec un locuteur du guarani (une langue parlée notamment au Paraguay), etc. : dans chacune de ces situations, nous sommes bien face à des langues différentes, et en effet on parle bien de la langue portugaise, de la langue hindi, de la langue danoise ou de la langue guarani. Inversement, quelqu'un qui parle le français, qu'il soit Bourguignon, Breton ou Parisien, n'a pas le sentiment de parler une langue différente, et de fait le français est considéré comme une seule et même langue. Mais les choses sont-elles toujours aussi nettes ?

Tout dépend en réalité de ce que l'on entend par « intercompréhension ». Un Français et un Italien peuvent parler chacun dans sa langue, et, moyennant quelques gestes, parvenir à avoir une conversation rudimentaire ; néanmoins, on ne saurait affirmer ici qu'il y a intercompréhension car tout Français et tout Italien sent bien qu'il parle une langue qui n'est pas la même que celle de son voisin d'une part, et ne peut le comprendre ni

être compris de lui couramment et sans effort d'autre part. Le français et l'italien sont donc bien deux langues.

Mais prenons maintenant le cas du bosnien, du croate, du monténégrin et du serbe. Nous sommes ici en face de quatre langues officielles. Pourtant, les différences entre ces quatre « langues » sont minimes, si bien que certains linguistes parlent de « BCMS ». C'est le cas de Paul-Louis Thomas et Vladimir Osipov, auteurs d'une *Grammaire du bosniaque-croate-monténégrin-serbe (BCMS)* (Thomas et Osipov, 2017). Voici à titre d'exemple l'article 1 de la Déclaration des droits de l'homme en bosnien, croate, monténégrin et serbe² :

Bosnien : Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i svijješću i treba da jedno prema drugome postupaju u duhu bratstva.

Croate : Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i svijješću i treba da jedno prema drugome postupaju u duhu bratstva.

Monténégrin : Sva ljudska bića rađaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i savješću i jedni prema drugima treba da postupaju u duhu bratstva.

Serbe : Sva ljudska bića radjaju se slobodna i jednaka u dostojanstvu i pravima. Ona su obdarena razumom i svešču i treba jedni prema drugima da postupaju u duhu bratstva.

Comme on le voit, ces quatre langues sont très similaires et il ne ressort aucune différence (sinon ponctuelle) des textes qui précèdent. De la même façon, le galicien et le portugais sont deux langues très proches l'une de l'autre ; pourtant, là aussi, elles sont, officiellement, deux langues différentes.

Inversement, certains dialectes d'Italie, comme le sicilien, n'ont pas d'intercompréhension nette avec l'italien. Néanmoins, le sicilien est considéré comme un dialecte, non comme une langue. Pour quelles raisons ?

Nous arrivons ici à un point important en ce qui concerne la définition de la langue, ou plutôt ses définitions possibles. Elles sont essentiellement au nombre de deux : linguistique et sociolinguistique.

D'un point de vue linguistique, on peut définir comme langue tout système de signes suffisamment différencié pour présenter un caractère qui lui est propre. Dans cette perspective, aussi bien le sicilien que le napolitain, par exemple, sont des langues. Le bosnien, le croate, le monténégrin et le serbe, par contre, sont une seule et même langue.

D'un point de vue sociolinguistique, l'idée de langue est associée à celle de prestige. Celui-ci peut être de plusieurs types : culturel, religieux, politique, de diffusion, etc. De ce point de vue, le bosnien, le croate, le monténégrin et le serbe sont bien quatre langues, associées à quatre États différents.

² Repris du site <<https://www.ohchr.org/EN/UDHR/Pages/Language.aspx?LangID=src2>> [consulté en mars 2020].

Qu'en est-il du tchèque et du slovaque ? On peut considérer qu'il s'agit de deux langues à la fois linguistiquement et sociolinguistiquement. Bien que proches et permettant l'intercompréhension, tout comme le norvégien et le suédois, chacune a en effet une physionomie qui la distingue de l'autre (critère linguistique) ainsi qu'un statut de langue officielle, à la fois dans leurs pays respectifs, la République tchèque et la Slovaquie, et dans l'Union européenne (critère sociolinguistique).

Comme nous l'avons observé, il s'avère difficile de donner une définition de la langue d'une part et des langues de l'autre, contrairement à ce que l'on pourrait penser au premier abord. Nous avons vu que la langue pouvait être définie comme un système de signes, qui sont utilisés oralement et éventuellement à l'écrit et reposent sur un principe de double articulation, d'après lequel un nombre fini de phonèmes permet de produire un nombre infini de signes. À cette définition interne, nous avons ajouté une définition externe, qui permet de délimiter les différentes langues parlées et aboutit à deux points de vue sur la langue. Selon un premier point de vue, qui est strictement linguistique, la langue est un système de signes qui est suffisamment caractérisé pour se distinguer d'un autre système de signes voisin. Selon un second point de vue, qui est sociolinguistique, une langue est un système de signes adopté par une communauté linguistique et promu à un certain prestige du fait de son caractère officiel, de la culture qu'il véhicule, de l'étendue de sa diffusion ou de toute autre raison.