

Univerzita Karlova
Filozofická fakulta
Ústav románských studií

Diplomová práce

PhDr. Bc.
Ondřej Drobník

Exploitation didactique des corpus dans l'enseignement du FLE à l'exemple des constructions verbonominales

Didaktické využití korpusů ve výuce francouzštiny na příkladu verbonominalních konstrukcí

Corpora in secondary FLE teaching:
case study on the verbonominal constructions

2022

Mgr. Magdalena Kučerová, Ph.D.

Shrnutí v češtině

Předkládaná magisterská práce se zabývá využitím česko-francouzského paralelního korpusu InterCorp ve středoškolské výuce. Teoretická část práce shrnuje formou vyjmenování několika výhod a nevýhod dosavadní poznatky korpusové didaktiky, zasazuje jazykový korpus do kontextu českého gymnaziálního vzdělávání a stručně popisuje konstrukci *avoir* N (např. *avoir faim*, *avoir besoin de*), jež bude jazykovým cílem našich cvičení. Metodologie představuje jednak data, se kterými se bude pracovat – paralelní korpus InterCorp (verze 13), jednak nastiňuje sestavení vlastního „didaktického“ korpusu, kterážto fáze je nevyhnutelná při čerpání dat z lingvisticky orientovaného korpusu. Těžiště diplomové práce spočívá v oddílu praktickém, kde učitelům franštiny navrhuje několik možností, jak sestavit vlastní cvičení pokrývající celý proces výuky cizího jazyka: expozici, fixaci, aplikaci i kontrolu (podle Choděry, 2013), přičemž se zaměřujeme na středoškolskou úroveň (cca A1–B1/B2). Navržená cvičení jsou dvého druhu: *hands-on* a *hands-off*. První typ pojednáváme pouze okrajově, u druhého jsme představili překladové, mluvní, mluvnické a j. aktivity. Hlavním výstupem práce je zjištění, že ve výuce lze paralelní korpus s úspěchem využít, a návrh několika způsobů, jak tento poznatek uplatnit v učitelské praxi.

Klíčová slova: FLE, paralelní korpus, gymnázium, *avoir* N

Résumé en français

Le présent mémoire porte sur l'exploitation du corpus parallèle tchéco-français InterCorp dans l'enseignement secondaire. La partie théorique résume les avantages et les inconvénients de la didactique de corpus, met le corpus dans le contexte de l'enseignement secondaire en République tchèque, et décrit brièvement la construction *avoir* N (*avoir faim*, *avoir besoin de*, etc.), qui servira de l'objectif langagier de nos exercices. La méthodologie présente le corpus parallèle InterCorp (version 13) qui sera exploité, ainsi que les grandes lignes de la construction de son propre corpus « didactique », phase inévitable lorsqu'on tire des données d'un corpus linguistique. L'accent du mémoire réside dans la partie pratique, où nous suggérons aux professeurs de français plusieurs manières de préparer leurs propres exercices couvrant l'ensemble du processus d'enseignement de langues étrangères : exposition, fixation, application et contrôle (selon Choděra, 2013), en nous focalisant sur l'enseignement secondaire (env. A1–B1/B2). Les exercices suggérés sont de deux types : *hands-on* et *hands-off*. Nous n'abordons le premier type que marginalement. Dans le cadre du second type, nous avons introduit quelques activités ciblant la traduction, l'oral, la grammaire, etc. Le résultat principal de ce travail est la constatation qu'un corpus parallèle peut être utilisé avec succès dans l'enseignement secondaire, et la proposition de plusieurs façons d'appliquer ces connaissances dans la pratique de l'enseignement.

Mots clé : FLE, corpus parallèle, lycée, *avoir* N

Abstract in English

The submitted master's thesis deals with the use of the Czech-French parallel corpus InterCorp in secondary school teaching. The theoretical part summarizes the already known advantages and disadvantages of corpus didactics, places the language corpus in the context of Czech high school education and briefly describes the construction of *avoir* N (e.g. *avoir faim*, *avoir besoin de*), which will serve as the language goal of our exercises. The methodology presents the data that will be worked with in the next parts, the parallel corpus InterCorp (version 13), as well as outlines the construction of own "didactic" corpus, which phase is inevitable when drawing data from a linguistically oriented corpus. The focus of the diploma thesis lies in the practical section, where I suggest to French teachers several ways to compile their own exercises covering the whole process of foreign language teaching: exposure, fixation, application, and control (according to Choděra, 2013), focusing on the secondary school level (approx. A1–B1/B2). The suggested exercises are of two types: hands-on and hands-off. I discuss the first type only marginally, in the second I have introduced translation, speaking, grammar and other activities. The main output of the work is the finding that a parallel corpus can be successfully used in secondary teaching, and the proposal of several ways to apply this knowledge in teaching practice.

Key words: FLE, parallel corpus, high school, *avoir* N

Za vedení, podnětné konzultace a pochopení mého ne přímo standardního způsobu organizace práce děkuju Mgr. Magdaleně Kučerové, Ph.D.

Prohlašuji, že jsem diplomovou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

PhDr. Bc. ONDŘEJ DROBNÍK

Table des matières

1	<u>INTRODUCTION</u>	<u>8</u>
2	<u>CORPUS DANS L'ENSEIGNEMENT DU FLE</u>	<u>11</u>
2.1	AVANTAGES	14
2.2	DESAVANTAGES	17
2.3	LE CORPUS DANS LE SYSTEME SCOLAIRE TCHEQUE.....	19
3	<u>CONSTRUCTIONS VERBO-NOMINALES</u>	<u>22</u>
3.1	THEORIE LINGUISTIQUE	22
3.2	LA CONSTRUCTION AVOIR N DANS L'APPRENTISSAGE DE LANGUE.....	24
4	<u>METHODOLOGIE</u>	<u>26</u>
4.1	SOURCE DES DONNEES	26
4.2	RECHERCHE ET TRI DES OCCURRENCES	28
5	<u>TYPES D'EXERCICES.....</u>	<u>34</u>
5.1	EXERCICES HANDS-OFF	35
5.1.1	ACTIVITE DE « MOTIVATION »	35
5.1.2	ANALYSE DE MINI-CORPUS	36
5.1.3	PHRASES A TROUS	40
5.1.4	DIALOGUES	41
5.1.5	DISCUSSION.....	41
5.1.6	REPERAGE DES TRAITS DU FRANÇAIS PARLE	42
5.1.7	TRADUCTION	42
5.1.8	ÉCRIRE UN TEXTE.....	43
5.1.9	REVISION D'UN PHENOMENE GRAMMATICAL.....	43
5.1.10	TEST.....	44
5.1.11	RESUME.....	45
5.2	EXERCICES HANDS-ON	47
5.2.1	PHRASEOLOGIE FRANÇAIS–TCHEQUE.....	48
5.2.2	PHRASEOLOGIE TCHEQUE–FRANÇAIS.....	49

5.2.3	RESUME.....	51
6	<u>CONCLUSION.....</u>	<u>53</u>
7	<u>BIBLIOGRAPHIE</u>	<u>55</u>
8	<u>APPENDIX.....</u>	<u>58</u>

1 Introduction

L'exploitation des corpus dans l'apprentissage des langues n'est rien de nouveau – au contraire, il y a déjà pas mal d'études faites dans ce domaine-là. Néanmoins, ces études ne portent, dans la plupart des cas, que sur l'enseignement de l'anglais et cela presque uniquement au niveau universitaire (cf. Boulton & Cobb, 2017 ; Šebesta & Šormová, 2019 ; Владиславович et al., 2020). De plus, ce type spécifique d'enseignement n'exploite, le plus souvent, que les corpus monolingues.

Le but de notre mémoire soit défini par rapport à ces trois prémisses générales. Dans ce travail, nous voulons traiter le français en tant que langue étrangère. Bien que les corpus aient déjà été exploités dans la didactique du FLE (p. ex. Di Vito, 2019) ou d'autres langues étrangères : cf. Forti (2019) pour l'italien, Vyatkina (2016) pour l'allemand, ou bien Šebesta & Škodová (2012) pour le tchèque, le nombre d'études et d'applications dans l'enseignement n'est pas si important que dans le cas de l'anglais.

Nous ne nous occuperons pas de l'enseignement universitaire, mais de l'enseignement secondaire (lycée, tch. *gymnázium*).¹ Ce niveau d'éducation a également ses traits spécifiques qui feront l'objet du travail. Ce domaine-là n'est pas, lui aussi, si bien décrit en comparaison avec l'enseignement universitaire (pour les rares études pratiques, cf. Moon & Oh, 2018).

Le troisième point dans lequel nous nous écarterons du courant principal est la nature de corpus. La plupart des didacticiens de corpus travaillent avec le corpus monolingue (*id est* un corpus qui ne contient qu'une langue). Pour notre projet, nous avons choisi le corpus parallèle (pour sa définition, voir le chapitre 2) qui, contenant deux ou plusieurs langues, l'une à côté des autres, a un potentiel didactique non négligeable. Dans ce mémoire, nous ciblons la découverte d'avantages de ce type de corpus.

Ainsi, nous avons préalablement défini le domaine de travail. Le point d'intersection des caractéristiques mentionnées ci-dessus contient les attributs suivants : 1. le corpus parallèle, 2. l'enseignement secondaire et 3. la langue française (alignée avec le tchèque – il s'agit de corpus parallèle). Dans le domaine tellement défini, nous n'avons pas trouvé de littérature – l'étude la plus proche de notre projet (respectant deux de nos trois critères) est celle de Nádvořníková et al. (2010) : *Využití InterCorpu ve vysokoškolských kurzech francouzské filologie*.

¹ Il s'agit, en général, du niveau de langue A1–B1/B2.

L'objectif du présent mémoire de Master est donc de traiter – et de remplir – les lacunes dans la didactique secondaire sur le corpus parallèle. Ce but veut toucher non seulement le repérage théorique, mais avant tout le côté pratique du problème. La présentation ciblera avant tout les enseignants qui, en s'inspirant de notre méthode, peuvent préparer leurs propres activités basées sur le corpus parallèle.

Les avantages et les inconvénients de l'emploi du corpus seront présentés au cours du mémoire (notamment dans les chapitres 2.1 et 2.2). En ce moment, on peut s'imaginer que le corpus parallèle, vu qu'il contient les traductions, serait utilisable pour présenter un élément compliqué (de point de vue de la langue) et relativement complexe. Plus concrètement, il serait possible d'envisager p. ex. la langue parlée, non formelle. Mais où sont les avantages ? Par exemple dans le fait que l'on peut consulter la traduction qui informe non seulement du « sens » de l'occurrence en question, mais aussi de sa valeur stylistique, communicative, etc. C'est pourquoi nous avons choisi le corpus parallèle en tant que source de données et pourquoi nous croyons favorable la présence de la langue maternelle dans l'enseignement d'une langue étrangère.

C'est pour cette raison que nous avons choisi la langue parlée en tant que « contenu » de nos futurs exercices. Étant donné que le présent mémoire ne veut aucunement décrire la langue parlée, elle ne sera présentée que brièvement, le seul objectif de cette présentation étant une vague esquisse de ce type de langue spécifique, notamment par rapport au contenu de notre corpus. Pour l'esquisser, nous nous servirons de l'ouvrage de C. Blanche-Benveniste (1997) intitulé *Approches de la langue parlée en français*, dont nous avons tiré les traits typiques² suivants du français parlé :

- question *in situ* : *Tu vas où ?*
- dislocation : *Mon père, il arrive.* (appelée parfois « redondance syntaxique »)
- 95 % d'absence de *ne* dans la conversation (i. e., uniquement l'« élément forclusif » est présent) : *Je sais pas.*
- interrogation par *est-ce que* : *Est-ce que tu veux ?*
- réductions : *T'as faim ?*
- question par l'intonation : *T'as faim ?*

² Nous avons restreint notre liste aux traits présents dans notre corpus. Cette liste nous sert « de référence » ; dans la suite du texte, une fois la langue parlée mentionnée, nous penserons toujours surtout à cette liste.

Après avoir fait cette réflexion, il faut se poser la question sur les données à traiter. Pour le moment, nous laissons la question non répondue parce que le choix et le tri des données est plutôt complexe et sera décrit en détail dans le chapitre 4. Néanmoins, nous pouvons esquisser la « clef » du choix de nos données : pour (dé)montrer l'un des avantages les plus frappants du corpus, nous avons choisi la construction *avoir* + NOM en tant qu'objet d'apprentissage de nos exercices (la construction sera présentée dans le chapitre 3). Cet avantage réside dans l'aisance avec laquelle le corpus cherche/trouve les occurrences de la construction choisie. Ainsi, on obtient un grand nombre d'occurrences pertinentes dans leur contexte (phrastique) « naturel ». Cela fait, à notre avis, la distinction majeure entre les concordances de corpus et une autre source authentique : tandis que le corpus offre pas mal d'occurrences dans un contexte plutôt restreint (dans les chapitres à venir, nous verrons pourquoi), la source authentique (p. ex. un texte tiré d'un journal) présente généralement peu d'occurrences dans leur contexte large.³

Pour conclure, notre objectif est de proposer quelques types d'exercices (voir le chapitre 5) basés sur le corpus qui respecteront les consignes suivantes – il s'agira d'exploitation de corpus parallèles, les activités cibleront l'enseignement secondaire, les langues incluses seront le français et le tchèque, et, pour ce qui concerne le contenu du corpus, nous traiterons la construction *avoir* + NOM dans le contexte de la langue parlée. La langue parlée ne jouera qu'un rôle secondaire, elle sera présentée aux étudiants « telle quelle », sans explications – sauf quelques exceptions. Pour ce qui concerne le public visé, nous écrivons ce travail avant tout de point de vue et pour les enseignants du FLE. Les exercices dans l'Appendix peuvent les inspirer et leur montrer comment on peut se servir du corpus pour préparer ses propres activités.⁴

³ En général, on pourrait parler du contexte « phrastique » pour le corpus, et du contexte « textuelle » pour le texte authentique. Dans ce travail, les actualisations seront présentées dans leur contexte phrastique – ce qui nous semble naturel, étant donné que le contexte « naturel » de la construction verbo-nominale *avoir* NOM est celui de la phrase.

⁴ On référencera nos activités, qui se trouvent jointes dans l'Appendix, toujours selon la matrice Ex. N, où N est pour le numéro de l'activité en question.

2 Corpus dans l'enseignement du FLE

Dans la partie préliminaire, nous avons plusieurs fois mentionné la notion de « corpus ». Bien qu'il ne soit pas notre but de définir ici le corpus de manière exacte et exhaustive, nous croyons bien d'indiquer au moins ce que le corpus (de langue) peut être, et quels sont les traits universaux des corpus. Par conséquent, on se passe de définition ciselée ; dans ce qui suit, nous présenterons la notion de « corpus » en s'appuyant sur les caractéristiques clef collectées par Čermák (2017b). D'après lui, le corpus est n'importe quel ensemble de textes aux traits suivants :

- **Vaste.** Si l'on se plonge jusqu'au fond de l'histoire de la linguistique de corpus, dans les années 1960, on se rend compte que les premiers corpus existants ne disposaient que de milliers, ou bien millions de mots au maximum (cf. p. ex. Čermák, 2017a). De nos jours, les corpus atteignent régulièrement le nombre de plusieurs millions ou milliards de mots textuels. Le corpus InterCorp, avec lequel nous travaillerons dans la suite, se classe parmi les corpus à taille moyenne, le sub-corpus franco-tchèque contenant à peu près 100 000 positions.⁵
- **Authentique.** Les corpus visent, en règle générale, à inclure les textes authentiques, c'est-à-dire les textes provenant d'une situation langagière concrète et réelle. Autrement dit, le corpus nous informe sur le fait comment la langue est réellement utilisée.
- **Écrit ou parlé.** Dans le cadre de ce travail, nous ne nous servirons que des corpus contenant les textes écrits. Néanmoins, il existe un nombre des corpus de textes parlés qui, eux aussi, possèdent un potentiel non négligeable dans le domaine de la didactique de langues.⁶
- **Électronique.** Le grand avantage, au moins par rapport aux travaux linguistiques dans le passé, est que les corpus sont entièrement mis en forme électronique. Ainsi, les corpus sont devenus très faciles à exploiter.

⁵ Le plus souvent, les linguistes de corpus parlent de « tokens » ou « mots textuels », autrement dit de mots graphiques. Ici, ce n'est pas le cas car dans les corpus sous la gestion de *Ústav Českého národního korpusu*, on ne compte pas seulement les « mots », mais aussi la ponctuation. Pour cette raison, on emploie le terme chapeautant de « position ».

⁶ Restent à côté les autres types de corpus : aujourd'hui, on a, de plus en plus souvent, affaire aux corpus multimodaux qui – outre le texte – contiennent des médias audio-visuels.

- **Mesurable.** Sous la mesurabilité, on comprend le fait que le corpus, en tant qu'ensemble fini de données langagières, peut devenir un objet de description statistique de n'importe quelle granularité et complexité. Ce qui est particulièrement important non seulement pour la linguistique en tant que science, mais aussi pour le présent mémoire (et nous le verrons plus tard), c'est la possibilité de mesurer la fréquence. La fréquence se fait valoir dans la lexicographie, la grammaire, etc. – et, ce qui est important, dans la didactique de langue. La raison est évidente : plus le mot s'emploie, plus grande doit être sa fréquence dans le corpus. Et *vice versa*, plus grande la fréquence, plus vraisemblable que l'on rencontre le mot dans l'usage.⁷ Comment nous avons travaillé avec la fréquence sera montré dans le chapitre méthodologique (cf. chapitre 4).
- **Non sélectif.** Les données stockées dans un corpus sont toujours traitées de manière non-sélective. C'est-à-dire que le concordancier ne nous cache aucune information demandée. La seule façon dont on peut influencer la sortie est le choix de données que l'on met à disposition. Ce point présente un problème délicat du point de vue didactique. À titre d'exemple, c'est dans le chapitre 4 que nous discuterons les problèmes liés à la présence de vulgarismes dans le corpus. Le concordancier, n'ayant, bien évidemment, pas d'entraves morales, n'est pas capable d'exclure les occurrences vulgaires.

Dans ce qui précède, nous avons présenté les traits centraux de (presque) tous les corpus. Étant donné que cette définition est assez large, on distingue plusieurs types de corpus. La distinction se fait à partir de différents critères : d'après le média, on distingue les corpus écrits, parlés, multimodaux, etc. ; selon la taille, les corpus se classifient comme petits, moyens, grands, etc. ; selon le but, on parle des corpus linguistiques, didactiques, etc. ; et ainsi de suite (pour une classification plus détaillée, voir p. ex. Chlumská, 2014). Une autre classification possible est celle qui s'intéresse aux langues comprises dans le corpus en question. Cette classification présente, vu le but du mémoire, la distinction clé. Si le corpus ne contient qu'une seule langue, on parle du corpus monolingue. *Per analogiam*, les corpus à plusieurs langues

⁷ Nous sommes conscients de la simplification que nous avons faite. Évidemment, on ne devrait pas oublier les facteurs qui peuvent remarquablement influencer la fréquence (p. ex. le contenu de corpus).

sont appelés multilingues. Parmi eux, on distingue p. ex. les corpus comparables, c'est-à-dire deux (ou plusieurs) corpus partageant une caractéristique – typiquement le contenu. Dans ce groupe de corpus appartiennent également les corpus parallèles qui présentent l'objet de notre travail. Un corpus parallèle est un corpus qui contient les mêmes textes dans deux ou plusieurs langues. Ces textes sont alignés les uns aux autres. Cet alignement a lieu aux différents niveaux – mais toujours, on aligne les unités de la même « granularité », de la même nature. Ainsi aligne-t-on p. ex. les paragraphes aux paragraphes, les phrases aux phrases, les mots aux mots, et ainsi de suite. Dans le cas du corpus InterCorp (exploité dans ce mémoire), nous aurons affaire à l'alignement au niveau phrastique.

En général, le corpus peut être considéré en tant que « produit ». Il ne faut pas oublier qu'il y a une étape précédente – la création du corpus (une description approfondie de tout le procédé de la création de corpus est à trouver p. ex. dans Pořízka, 2014). Ce thème a fait couler déjà beaucoup d'encre, c'est pourquoi nous nous restreignons à une seule remarque : le corpus se crée toujours avec certaine intention, autrement dit, sa création poursuit un but concret. Le plus souvent, ce but est d'ordre linguistique – rares sont les corpus didactiques,⁸ d'autant plus pour nos deux langues, sans mentionner la paire français-tchèque. Le corpus InterCorp a été également créé avec une intention linguistique.

Il nous semble évident que, dans l'enseignement, chaque type de corpus peut être employé d'une manière appropriée, bien que différente. Dans le milieu universitaire, les corpus monolingues s'emploient notamment pour vérification d'hypothèses, tandis que les corpus parallèles sont utilisés pour la recherche de similarités et contrastes (cf. Nádvorníková et al., 2010). Vu que l'application de corpus dans les écoles secondaires n'est pas si bien décrite, nous n'avons pas réussi à trouver une mention de l'emploi d'autres corpus que monolingues.

Dans le reste du chapitre, nous définirons les termes centraux, qui sont traditionnellement liés à la linguistique de corpus, et esquisserons leurs significations dans ce travail.

Le premier ensemble de termes à expliquer est la triade *corpus-based*, *corpus-driven* et *corpus-informed* (cf. aussi Čermák, 2017a). Nous laissons de côté les discussions linguistiques et tenons les approches *corpus-based* et *corpus-informed* pour synonymes. Ces

⁸ Cette distinction concerne toujours le but : le corpus linguistique a été créé « pour les linguistes », afin de décrire la langue, tandis que le corpus didactique se crée « pour les enseignants », p. ex. pour préparer les exercices. Ainsi, une différence réside également dans le contenu : pour les linguistes, il est important d'explorer la langue « telle quelle » ; les didacticiens veulent, au contraire, que les données soient ajustées au niveau d'étudiants. Donc, on peut dire que le corpus didactique ne partage pas, *stricto sensu*, la « non-sélectivité » (voir *supra*).

termes veulent dire « se servant du matériel provenant de corpus ». Le troisième terme, *corpus-driven*, veut dire « travaillant sans hypothèses posées *a priori* ». Autrement dit, dans ce cas ce sont les données qui dirigent la recherche.

Une autre paire terminologique important pour nos buts est *hands-on* et *hands-off*. Dans le domaine de la linguistique de corpus appliquée, on emploie le terme *hands-on* (p. ex. « exercices *hands-on* ») pour les activités pendant lesquelles les étudiants travaillent directement avec le corpus. *Hands-off* veut dire que les étudiants ne se voient pas exposés aux corpus – les concordances sont choisies en avant par l’enseignant et, le plus souvent, remaniées d’une certaine manière et éventuellement imprimées.

Finalement, nous tenons à une brève définition de l’enseignement du FLE. Sous le terme « enseignement », nous comprendrons, dans le contexte de ce travail, l’enseignement secondaire (étudiants entre 15–19 ans). Sous le terme « FLE », on comprend le « français langue étrangère » ; donc les étudiants non natifs de la langue français – ici, nous visons notamment les locuteurs natifs tchèques. Cette délimitation est assez générale, elle sera ciselée en passant dans les chapitres à venir.

2.1 Avantages

Dans les deux chapitres suivants, nous résumerons – en nous appuyant sur la littérature scientifique – les avantages et les inconvénients de l’application de corpus dans l’enseignement de langues étrangères. Comme déjà prévu, la littérature du domaine en question s’intéresse notamment au niveau universitaire (laissant de côté le fait que la plupart des études a été menée auprès des apprenants de l’anglais). Néanmoins, nous considérons – et certains auteurs le disent de manière explicite (p. ex. Chun, 2016) – que les phénomènes présentés ci-dessous sont plus généraux. Ainsi, nous traiterons les points les plus importants qui, en surcroît, se sont avérés vrais même en pratique, lors de notre pilotage.

Commençons par une redite de la définition du corpus proposée dans le chapitre précédant (chap. 2). Nous verrons que certains avantages sont à chercher dans la nature même du corpus langagier. Il s’agit notamment du fait que les données dans le corpus sont authentiques et non sélectives. Par rapport à la didactique, il devient évident que, en exploitant le corpus, on ne présente pas d’exemples forgés ; tout au contraire, on a toujours affaire à la langue « telle qu’elle est » (plus en détail traité dans Kamber & Dubois, 2016). Ce qui s’avère également relativement novateur,⁹ notamment en comparaison avec certains manuels datés, est

⁹ Au cours de l’histoire de la didactique de langues étrangères, les textes authentiques n’étaient pas toujours préférés. Un tournant vient avec les méthodes plutôt modernes, p. ex. avec l’approche communicative.

la possibilité de présenter la langue tout à fait actuelle et, en même temps, dans ses divers registres, « niveaux ».¹⁰

De plus, ces occurrences, authentiques et non sélectives, sont contextualisées (Günday et al., 2017). C'est-à-dire que l'on peut voir les contextes gauche et droit, et les employer dans l'enseignement. Il faut distinguer aussi les deux « axes » du corpus, ou plutôt de la langue en tant que système (cf. p. ex. Saussure, 1996) : l'axe syntagmatique (relations *in praesentia*) et l'axe paradigmatique (relations *in absentia*). De point de vue didactique, il faut que les étudiants apprennent à observer et à généraliser les deux.

Par cette observation, nous parvenons à l'une des caractéristiques les plus pertinentes de l'apprentissage sur corpus : il s'agit d'une procédure inductive. Autrement dit, les étudiants sont, le plus souvent, invités à travailler avec les données authentiques, à en tirer eux-mêmes « leurs » règles grammaticales, etc. Tout cela, nous le rappelons, en ayant les deux axes langagiers devant les yeux.

Ainsi, les étudiants deviennent « détectives de la langue » (*language detectives*, terme de T. Johns). Par ce type d'apprentissage, on soutient et développe aussi les capacités cognitives, la pensée critique et les *problems-solving skills* (d'après Chun, 2016). Vu les exigences du système scolaire tchèque (plus en détail dans le chapitre 2.3), le développement des « compétences » est vu très favorable dans l'enseignement de langues étrangères. En effet, utilisant le corpus, on n'enseigne pas seulement la langue, mais aussi ces capacités et, avant tout, l'approche critique de données et d'informations ce qui s'avère, à nos jours, une capacité nécessaire.

Pour être plus concret, nous pouvons citer un exemple de ce qui peut être (systématiquement) développé en travaillant avec les corpus : si l'enseignant mène ses étudiants vers la faculté de bien reconnaître les structures de la langue, l'activité durant laquelle les occurrences tirées du corpus se montrent un outil idéal, il est vraisemblable qu'ils apprennent à identifier ces structures plus facilement. Donc, ils apprennent la *colligational awareness* (Flowerdew, 2009 ; « sensibilité aux cooccurrences catégoriques » – traduit par l'auteur). Cette faculté concerne p. ex. les verbes à rection prépositionnelle (« préposition vide ») ou bien la construction décrite dans le cadre du présent mémoire : *avoir* à complément nominal (*avoir* N).

Pourtant, cette « awareness » ne concerne pas uniquement le côté formel du signe, mais aussi la composante sémantique, « fonctionnelle ». Ces deux côtés du signe sont, d'une certaine manière, liés aux axes de la langue décrits ci-dessus. Tout comme p. ex. la

¹⁰ Bien sûr, les « outputs » sont, comme nous l'avons déjà montré, toujours tributaires des « inputs »...

syntagmatique, la sémantique, elle aussi, se laisse reconstruire à partir des occurrences de corpus. À titre d'exemple, prenons de nouveau la construction *avoir* N, concrètement *avoir chaud*. Comme nous le verrons dans les exercices (voir chapitre 5 ou l'Appendix), cette locution peut avoir deux sens : le sens compositionnel (antonyme à *avoir froid*) et le sens non compositionnel, figuré (égal à *il s'en est fallu d'un cheveu*). À l'aide des données du corpus parallèle, les étudiants apprennent sans fatigue les deux sens et, de plus, l'emploi approprié.

Dans le chapitre 2 nous avons dit que le corpus est mesurable. Sa mesurabilité peut être profitablement appliquée dans la didactique. Le plus souvent, on se sert de la fréquence : à partir du corpus, on peut facilement dresser une liste de fréquence qui nous dit p. ex. quels sont les mots les plus utilisés dans une langue concrète. Ainsi, les enseignants peuvent travailler avec les mots les plus courants, les auteurs de manuels peuvent planifier l'exposition du lexique correspondant au niveau concret, etc. Dans ce travail, et nous le verrons plus tard (voir chapitre 4), l'information sur la fréquence des éléments différents a été également prise en considération.

Parallèlement aux facultés indiquées jusqu'ici, l'application de corpus en classe peut aussi soutenir les *digital skills* (Aubin & Bírová, 2019 ; « compétences numériques »). Si l'on exploite les corpus d'une manière *hands-off*, il faut comprendre cet avantage d'une manière plus étroite. Les apprenants développent la faculté d'induire, de généraliser, d'approcher les données d'une façon critique, etc. Pour l'approche *hands-on*, on développe, à côté desdites, la capacité de manier l'ordinateur et, plus spécifiquement un outil concret, le corpus. Ce point compte, à notre avis, notamment dans le milieu universitaire, linguistique. Là, les étudiants sont souvent censés faire leurs propres recherches sur les corpus. Néanmoins, nous croyons que le corpus peut être introduit, en tant qu'outil didactique, même dans les autres types d'écoles. C'est déjà avec une connaissance infime que l'on peut consulter le corpus en en profitant : les apprenants peuvent l'utiliser p. ex. en écrivant un travail de séminaire, pour y vérifier les constructions langagières (cela se montre valable et favorable même pour les locuteurs natifs – cf. André, 2020).

Le dernier point que nous traiterons dans ce chapitre est la motivation. La motivation des apprenants est un terme fréquemment décliné dans le discours didactique (cf. p. ex. Choděra, 2013 ; Podlahová, 2012 ; Skalková, 2007) qui, bien évidemment, touche également le domaine de l'application de corpus dans l'enseignement. Dans les ouvrages scientifiques (p. ex. Anderman & Anderman, 2021), la motivation paraît une question omniprésente. Pour motiver les apprenants à apprendre la langue à l'aide des corpus de langue, on mentionne plusieurs éléments motivants. Par exemple, Nádvorníková et al. (2010) identifient parmi les

sources de motivation le travail en groupe. À côté de la collaboration (en classe ou entre les classes), les scientifiques parlent également du fait que les corpus présentent une approche tout à fait moderne et nouvelle pour les apprenants ce qui, au moins au début du processus de l'apprentissage, peut être considéré comme une autre source de motivation pour l'apprentissage. Dans les études déjà faites (cf. *inter alia* Chun, 2016 ; Flowerdew, 2009), les auteurs problématisent aussi le contact (souvent direct, notamment dans le milieu universitaire) avec l'ordinateur qui fait éveiller le « plaisir de recherche », étant donné que l'approche donnée suscite la « curiosité naturelle » des apprenants.

Jusqu'ici, nous avons parlé de la motivation des étudiants pour l'apprentissage de la langue – en nous appuyant sur la littérature existante sur ce thème. Néanmoins, à ce que nous savons, personne n'a systématiquement traité la motivation des enseignants pour l'intégration du corpus dans les cours de langues.¹¹ Ici, nous laissons la question ouverte – elle sera élaborée plus en détail dans la section 6.

2.2 Désavantages

Pour rester objectif et pour pouvoir travailler de manière appropriée, il ne nous suffit pas d'avoir mentionné les avantages. Il faut rappeler, à côté des avantages, également les désavantages. Dans le présent chapitre, nous résumerons donc les inconvénients mentionnés dans la littérature.

L'application de corpus dans l'enseignement, bien que pratiquée déjà à partir des années 70 du siècle précédent (cf. p. ex. McEnery & Wilson, 2001), a toujours, même un demi-siècle après sa naissance, l'air d'une méthode nouvelle, non « enracinée ». Certes, on se sert des corpus dans les domaines de lexicographie, grammatographie, etc. – les apprenants sont donc, à la fin, réellement exposés aux données du corpus ; mais le travail plus « explicite », focalisé sur les données (voire le travail *hands-on*) reste toujours limité à quelques centres universitaires. Ainsi, le corpus demeure pratiquement inconnu parmi les enseignants de lycée.¹²

Un autre inconvénient pour cette recherche pourrait être désigné comme manque relatif de littérature scientifique. Comme nous l'avons déjà plusieurs fois indiqué, la littérature sur le thème se restreint sur le domaine universitaire. Pour l'éducation secondaire,

¹¹ C'est seulement Boulton (Boulton, 2017 ; Boulton & Cobb, 2017 ; Leńko-Szymańska, 2017) qui a élaboré la question du point de vue de professeurs.

¹² D'une manière non officielle et non systématique, nous nous sommes informé sur ce que les enseignants (niveau lycée) savent de l'emploi du corpus. Le plus souvent, ils ont déclaré qu'ils n'avaient jamais rencontré ce terme-là... La cause de cette situation peut être, à notre avis, double : premièrement, certains enseignants ont fini leurs études avant l'expansion de corpus, donc ils ne le connaissent pas de l'école, deuxièmement, on ne propage pas évidemment le corpus dans ce type d'enseignement.

nous n'avons trouvé que quelques études. Néanmoins, nous sommes convaincu que la différence entre les lycées et le niveau universitaire n'est pas si frappante, comme nous le verrons plus tard (cf. le chapitre 5). À notre avis, la seule différence qu'il faut prendre en considération, est, d'après notre expérience didactique parallèle dans les deux domaines, le degré de connaissances professionnelles (c'est-à-dire linguistiques).¹³ De cette constatation découle la nécessité d'ajuster la forme des matériaux employés dans l'enseignement (il faut p. ex. limiter les occurrences de termes scientifiques, adapter le travail de « détective de langue », etc.).

Les auteurs, parmi autres Nádvorníková et al. (2010), mentionnent que la capacité de travailler avec les corpus est indispensable pour leur emploi dans l'enseignement, autrement dit qu'il faut posséder les connaissances pratiques avant de pouvoir apprendre à partir des corpus. Néanmoins, cette affirmation a deux « points faibles » de point de vue de notre projet : le premier concerne le simple fait que ces études ont été faites dans le milieu universitaire, le plus souvent linguistique. Par conséquent, il est évident que les linguistes doivent, pour en pouvoir profiter, bien maîtriser les corpus.¹⁴ La seconde remarque qu'il faut faire par rapport à la littérature citée est la suivante : les mentions sur la « technicité » et certain degré d'inaccessibilité des corpus sont, le plus souvent, relativement datées, voire dépassées. La raison est simple – dans ces dernières années, les corpus ont beaucoup développé, et, d'une manière outrée, on peut dire que la recherche (élémentaire !) dans le corpus ressemble considérablement aux recherches sur les moteurs de recherche, comme p. ex. Google.¹⁵ Bien sûr, on ne fait pas de *rocket science* en se servant uniquement de la recherche élémentaire, mais nous sommes pourtant convaincu qu'elle peut devenir un complément attractif du processus d'apprentissage. Les conséquences pratiques seront montrées et discutées dans les parties 5 et 6.

Les autres problèmes touchent l'organisation « pratique » du travail (non seulement en classe). Dans le cadre de l'enseignement secondaire, les enseignants se posent souvent la question sur les finances. Il existe une vaste quantité de corpus dont certains sont payés, certains non. Le côté « pratique » est donc à examiner avant la phase didactique.

Un autre désavantage concerne la nécessité d'assistance pédagogique au cours d'apprentissage. L'assistance est souvent discutée dans le contexte des exercices *hands-on*. D'après certains auteurs, ces exercices doivent être guidés par le pédagogue (Di Vito, 2013 ;

¹³ Nous comparons ici les lycées généraux aux étudiants en linguistique. À part restent les étudiants d'autres domaines, bien que cette problématique – de nouveau au niveau universitaire – ait été relativement profondément décrite.

¹⁴ Cette remarque ne concerne uniquement le corpus « en tant que tel », mais aussi p. ex. la description statistique adéquate, etc.

¹⁵ Il faut se rendre compte que différents corpus utilisent différents concordanciers. Ici, nous nous concentrons notamment sur les corpus sous la gérance de *Ústav Českého národního korpusu* (voir <https://korpus.cz>).

Di Vito, 2019) afin de « protéger » les apprenants d'un afflux excessif de données. Sinon, les étudiants ne sauraient pas s'y prendre de manière efficace, et tout cela mènerait à leur démotivation – dans les sources scientifiques, on appelle cette perte de motivation le *learning fatigue* (Chun, 2016).

Très étroitement liée avec ce que nous avons abordé ci-dessus est la nature chronopage (Chun, 2016) du travail avec les corpus – de nouveau notamment dans le cas des exercices *hands-on*. La raison, que nous voyons, semble multiple : il s'agit le plus souvent des activités inductives ; les apprenants ne maîtrisent pas les corpus ; parfois, on a affaire aux concordances linguistiquement assez compliquées, bien qu'intéressantes ; et ainsi de suite...

Le dernier des inconvénients que nous rappellerons, est le manque des corpus spécialisées, notamment pour nos deux langues – sans mentionner la nature « parallèle » du présent projet. Pour l'anglais, on a déjà dressé quelques corpus à but didactique (pour les exemples, nous référons à Šormová & Šebesta, 2019) – en ce que nous savons, pour le français aussi bien que pour le tchèque, ce type de corpus n'existe pas (encore). Dans ce travail, nous nous sommes donc restreint aux corpus linguistiques. La conséquence la plus importante de notre décision est le tri nécessaire des concordances dont le but est l'élimination des données inappropriées pour l'application didactique.

2.3 Le corpus dans le système scolaire tchèque

Dans cette partie, nous esquisserons tout brièvement l'ancrage de notre mémoire dans le concept de l'éducation nationale en République tchèque. C'est ici que nous voulons montrer que l'emploi du corpus en classe ne s'oppose pas à ce concept ; exactement le contraire est vrai. Le travail avec le corpus (sous n'importe quelle forme) peut développer plusieurs domaines de compétences visées par le programme national. Ce chapitre portera sur deux thèmes : *primo*, ce seront les rapports du corpus avec les « compétences », *secundo* la relation avec l'enseignement de la « deuxième langue étrangère », tel qu'il est présenté dans le programme cadre : *Rámcový vzdělávací program pro gymnázia* (« RVP » dans la suite du texte).

Pour ce qui concerne les « compétences clé » (tch. *klíčová kompetence*), l'exploitation de corpus, telle que conçue dans ce mémoire, en développe au moins trois (voir RVP, p. 9) :

1. compétence d'apprendre (tch. *kompetence k učení*)
2. compétence de résoudre des problèmes (tch. *kompetence k řešení problémů*)
3. compétence communicationnelle (tch. *kompetence komunikativní*)

La première compétence se voit développée par la conception de l'enseignement – on accentue le travail autonome et les styles d'apprentissage d'une langue étrangère à partir de données réelles. La capacité de résoudre les problèmes est soutenue notamment par l'accent mis sur les analyses de la langue (autrement dit, par l'approche inductive). La compétence de communication est développée avant tout en intégrant les connaissances « langagières » dans le contexte réel, authentique.

Vu que le corpus se base *per definitionem* (voir sa définition dans le chapitre 2) sur la technologie, le travail avec le corpus demande, en effet, un travail avec les « nouvelles technologies » (bien qu'il soit direct ou indirect). On en déduit que, en exploitant le corpus, on ne développe pas seulement les capacités mentionnées ci-dessus, mais aussi la « compétence digitale » (tch. *digitální gramotnost*). Ainsi, notre procédé est parfaitement compatible avec le « concept de développement des compétences digitales » (tch. *koncept rozvoje digitální gramotnosti a informatického myšlení dětí a žáků*).¹⁶ Dans ce concept, on mentionne que les nouvelles technologies peuvent être intégrées dans l'enseignement de n'importe quelle matière (ou, plus exactement, de toutes les matières). De plus, le concept cité distingue deux catégories : la « compétence digitale » (déjà mentionnée) et la « pensée informatique » (tch. *informatické myšlení*). La seconde catégorie a plusieurs sous-catégories, dont au moins deux sont – selon les descripteurs proposés – développées en exploitant le corpus : premièrement, c'est le traitement critique des données (tch. *získávat, zaznamenávat, uspořádávat, strukturovat, předávat data a informace*), deuxièmement la capacité de trouver une/la solution¹⁷ (tch. *testovat, analyzovat, vyhodnocovat, porovnávat a vylepšovat uvažovaná řešení*).

En parlant du rapport entre le RVP, le FLE et le corpus, nous nous appuyerons sur le chapitre *Další cizí jazyk* du RVP (p. 19ff ; « deuxième langue étrangère »). Cette partie du RVP décrit, *grosso modo*, les buts de l'enseignement dans le domaine de la deuxième langue étrangère. Le niveau langagier cible se décrit comme suit : *Vzdělávání v oboru Další cizí jazyk navazuje na úroveň jazykových znalostí a komunikačních dovedností odpovídajících úrovni A1 [...] a směřuje k dosažení úrovně B1 [...]* (RVP, p. 21). Dans nos exercices proposés dans le chapitre 5, nous ciblons exactement ce niveau-là, donc A1–B1.

¹⁶ La description du concept est disponible sous <https://www.nuv.cz/t/koncept-rozvoje-digitalni-gramotnosti-a-informatickeho> (cité le 7 avril 2022) ; le site est la source de ce qui suit. D'ailleurs, le site propose une définition de la « compétence digitale » : *Digitální gramotností rozumíme soubor digitálních kompetencí (vědomostí, dovedností, postojů, hodnot), které jedinec potřebuje k bezpečnému, sebejistému, kritickému a tvořivému využívání digitálních technologií při práci, při učení, ve volném čase i při svém zapojení do společenského života.*

¹⁷ À titre d'exemple, on peut citer une activité où les étudiants sont invités à « découvrir » la règle grammaticale.

Pour ce qui concerne le développement des compétences langagières spécifiques, notre projet vise, entre autres, les deux capacités suivantes : 1. [žák] *rozliší v mluveném projevu jednotlivé mluvčí, identifikuje různé styly a citová zabarvení promluvy* (« l'élève est capable de distinguer les locuteurs, les registres ou le ton affectif ») et 2. *odhadne význam neznámých slov na základě již osvojené slovní zásoby a kontextu* (RVP, p. 19 ; « l'élève devine le sens des mots nouveaux en s'appuyant au contexte et au vocabulaire qu'il connaît » – les deux traductions sont nôtres). C'est avant tout le point 2 qui est développé par les activités basées sur les données du corpus (langue authentique, traductions non littérales, etc.).

3 Constructions verbo-nominales

3.1 Théorie linguistique

Avant d'avancer vers les exercices mêmes, leur préparation, les sources de données, etc., nous présentons le phénomène grammatical choisi, qui nous accompagnera dans tout le mémoire. Nous avons décidé de travailler sur les constructions verbo-nominales avec une structure formellement bien définie : nous nous intéresserons aux locutions avec le pivot *avoir* et un complément (« direct ») nominal sans détermination (dans le texte suivant, cette construction sera marquée par *avoir* N). Ce manque de déterminants peut théoriquement, et nous le verrons ci-dessous, indiquer un certain degré de lexicalisation, voire « phraséologisation ».

Notre recherche sur la construction *avoir* N est double : d'un côté, nous avons effectué la recherche dans les grammaires, de l'autre dans les dictionnaires, étant donné que cette construction se trouve à la frontière de la grammaire et du lexique. Vu l'orientation entièrement contrastive du projet, nous avons consulté avant tout les ouvrages contrastifs.

Les grammaires, en général, classifient la construction *avoir* N comme une « locution verbale » (en tchèque « slovesné vazby ») ; ainsi p. ex. Hendrich et al. (2001, p. 440), Šabršula (1986, p. 95), mais aussi Grevisse & Goosse (2016, p. 209), Riegel et al. (2018, p. 312), ou bien une monographie de J. Radimský (2010). *Le Bon Usage* propose, en supplément, une définition générale de « locution » – d'après eux, c'est une *suite de mots [...] qui forment [...] une unité lexicale* (Grevisse & Goosse, 2016, p. 209). Cela se fait remarquer également par équivalence à un mot simple – donc, par exemple, la locution *avoir lieu* correspond au verbe dit « simple » *arriver*.

Ces locutions verbales, comme déjà présenté ci-dessus sous le prisme de la grammaire de construction, se voient composées d'un « verbe support » (Riegel et al., 2018, p. 415)¹⁸ et d'un substantif jouant le rôle de « lexème factif » (Šabršula, 1986, p. 95). Le substantif figure, dans la construction, le plus souvent sans détermination – J. Šabršula (*ibidem*) considère que *člen není ve verbonominálních komplexních semionech staršího původu* (« l'article est absent dans les vieux sémiones complexes verbo-nominaux » – traduit par l'auteur). Néanmoins, le déterminant (l'article) peut apparaître dans le cas où le substantif est élargi, p. ex. par épithète ou complément (ainsi, la phrase *avoir une faim de bête*¹⁹ est jugée parfaitement grammaticale). Une autre modification possible se fait par adverbe « interposé »

¹⁸ Outre *avoir*, ils citent aussi les verbes *faire*, *rendre*, etc.

¹⁹ L'exemple repris de Šabršula (1986, p. 96).

– Grevisse & Goosse (2016, p. 209) citent un exemple typique pour le français parlé : *Il a très faim*. Voilà une particularité syntaxique, à savoir un substantif modifié par un adverbe à sens scalaire. À l'ordinaire, cette modification n'est pas jugée grammaticale : * *Il a très un frère*. Cette caractéristique syntaxique pourrait figurer comme un argument pour un statut spécial de ces constructions.

Les constructions *avoir* N se trouvent à la frontière de la grammaire et du lexique du français. Elles ne sont ni complètement « aléatoires », libres, ni complètement lexicalisées. Autrement dit, on a affaire au groupe de collocations très bien établies dans la langue française qui, au contraire, ne se comportent pas de manière tout à fait « lexicale » (cf. p. ex. l'insertion d'adverbe). C'est pour cela qu'elles sont jugées *lexicalisées ou quasi lexicalisées* (Riegel et al., 2018, p. 312). Grevisse & Goosse (2016, p. 209) ajoutent que bien que les constructions *avoir* N soient lexicalisées (ou « quasi lexicalisées » ; voir *supra*), les locuteurs sont capables de voir leurs composantes. Ainsi, on peut mettre en opposition p. ex. *avoir soif* et *avoir faim*, où les noms *soif* et *faim* sont paradigmatiquement interchangeables.

En ce qui concerne les constructions « remplies », les grammaires (toutes citées ci-dessus) mentionnent les locutions suivantes (rangées par ordre alphabétique) : *avoir besoin*, *avoir droit*, *avoir envie*, *avoir peur*, *avoir soif*. Cette liste correspond, à peu près, à la nôtre dressée dans le chapitre 4.2.

Deuxièmement, nous avons consulté les dictionnaires, notamment bilingues. D'abord, nous avons cherché le lemme *avoir* auprès duquel nous nous sommes attendu aux explications portant sur l'emploi phraséologique/collocationnel. Nous n'avons pas trouvé de mentions dans les dictionnaires imprimés (p. ex. dans Vlasák, 2002). Néanmoins, les dictionnaires en ligne²⁰ surabondent en entrées pertinentes – nous ajouterons les traductions trouvées à la liste présentée dans le chapitre 4.2.

Les raisons pour le choix de cette construction sont les suivantes : premièrement, comme nous l'avons déjà indiqué, cette construction est, de point de vue formelle, facile à définir.²¹ Deuxièmement, même après n'avoir fait que quelques recherches préparatoires dans les grammaires, nous nous sommes rendu compte de la variabilité de *avoir* N du point de vue didactique – les constructions que l'on présente dans les grammaires couvrent une échelle relativement large des niveaux de connaissance de la langue. Pour être plus concret, la locution

²⁰ Nous nous servons de deux dictionnaires tchèques en ligne : *Seznam slovník* (sous <https://slovník.seznam.cz>) et *Slovník Lingea* (sous <https://slovníky.lingea.cz/francouzsko-cesky>).

²¹ Ce qui est un trait très avantageux en travaillant avec le corpus, où la « communication » entre l'utilisateur et l'ordinateur se base sur les « formalismes »...

avoir faim ne se range, vu sous ce prisme, au même niveau que, par exemple, le phraséologisme *On a eu chaud*. Et troisièmement, ce phénomène grammatical est très riche en formes – si on le tenait pour une *construction* (pris au sens strict de la grammaire de construction, cf. p. ex. Fried & Östman, 2004, ou Mathieu, 2003), il s’agirait d’une construction semi-schématique à deux composantes : la première étant le verbe *avoir* sous une de ses formes fléchies, la seconde un complément défini par les valeurs « nominal » et « sans détermination ». Ainsi, cette « vacance » peut être remplie par une quantité non négligeable de noms : *faim, soif, besoin...* Nous reviendrons à ces caractéristiques dans les parties à venir.

3.2 La construction *avoir N* dans l’apprentissage de langue

Le fait que la construction *avoir N* varie selon les « actualisations » de sa partie nominale, a ses conséquences de point de vue de la position dans le parcours de l’apprentissage de la langue, c’est-à-dire de point de vue des niveaux de connaissances. Pour être plus concret : par exemple, la locution *avoir besoin de QC* ne partage pas, vu sous ce prisme, le même statut avec la locution idiomatique *on a eu chaud* – tandis que la première se trouve tout à la base des connaissances du français (niveau A2, à peu près), la seconde la surmonte considérablement : vu son sens non-compositionnel, le style/contexte restreint d’usage (*i. e.* français parlé, familial), la seconde locution ne correspond pas au niveau élémentaire ; tout au contraire, il s’agirait, d’après les descripteurs du « Cadre européen commun de référence pour les langues » (CCRL), plutôt d’un niveau plus élevé, disons B2/C1.²²

C’est pour cette raison qu’il faut trier les occurrences et trouver les plus convenantes. D’après le RVP (p. 13), nous le rappelons, l’éducation dans le cadre de la « deuxième langue étrangère » (tch. *další cizí jazyk*) commence par le niveau A1 et, après ces quatre années, les étudiants sont censés d’atteindre le niveau B1. Ainsi, nous choisirons les locutions correspondant à ces niveaux-là.

Pour ce qui concerne le « plan » langagier qui sera développé, nous en ciblons plusieurs. Selon ce que propose le RVP (p. 20 ; il s’agit du paragraphe *Žák se naučí*, « l’élève apprendra »), notre thème est le plus proche, nous le croyons, de la lexicologie. Ici, l’apprenant doit apprendre les collocations et locutions figées (tch. *ustálená slovní spojení*). Néanmoins, notre conception du thème dépasse le champ lexicologique, il touche également la phonétique, notamment en entraînant la prononciation des paires minimales ou des « difficultés » (RVP,

²² Dans ce chapitre, nous nous servons des descripteurs et, plus en général, de la conception du CCRL présentés dans la « version actuelle » du CCRL, décrite sous <https://www.msmt.cz/mezinarodni-vztahy/spolecny-evropsky-referencni-ramec-pro-jazyky> (cité le 9 avril 2022). Nos estimations s’appuient sur les descripteurs du CCRL.

p. 20 ; tch. *lexikologie – ustálená slovní spojení* ; pour l'application didactique, voir le chapitre 5.1.9), ou le domaine de la grammaire. Pour la grammaire (dans le RVP, p. 20, sous l'étiquette *gramatika*), on touchera, dans la partie pratique (voir chapitre 5), à titre d'exemple :

- les conjugaisons verbales (surtout formes fléchies du verbe *avoir* ; RVP : *jednoduché slovní tvary*),
- la structuration du lexique (p. ex. les relations antonymiques ; RVP : *antonyma*),
- la révision des temps verbaux (RVP : *základní vyjádření přítomnosti, minulosti a budoucnosti*),
- les locutions figées (RVP : *slova složená a sousloví*).²³

Ce chapitre a traité, tout en général, la position de la construction *avoir* N dans l'apprentissage du FLE. Ce n'est, pourtant, que le chapitre 4.2 qui présentera les locutions concrètes sur lesquelles nos exercices seront fondées. Là, on proposera la liste des locutions qui sera accompagnée non seulement de traductions, mais aussi d'indication approximative du niveau langagier correspondant.

²³ *Nota bene* : bien que la distinction entre *ustálené slovní spojení* (lexicologie) et *slovo složené a sousloví* (grammaire), proposée par le RVP (p. 20), ne nous soit pas claire, ce « chevauchement terminologique » confirme, quoiqu'indirectement, que la construction *avoir* N appartient aux deux domaines.

4 Méthodologie

Le but de cette partie méthodologique est double : *primo*, nous présenterons la démarche que nous avons adaptée en cherchant les données, en choisissant les occurrences pertinentes et en créant les exercices de langue. *Secundo*, nous voulons discuter également certains problèmes liés au tri de données langagières à but didactique. En effet, ce chapitre, exactement comme le fera la partie suivante, ne nous servira pas seulement d'introduction d'une méthode, mais aussi de résumé des obstacles qu'il faut prendre en conscience en exploitant les corpus parallèles dans les cours de langue étrangère.

4.1 Source des données

Tout d'abord, nous présenterons le (sub)corpus avec lequel nous avons travaillé. Comme nous l'avons mentionné plus haut, le but du mémoire est de proposer la méthodologie d'exploitation de corpus parallèles dans l'enseignement secondaire. Notre intention est de thématiser notamment la langue parlée. Cela est la raison pour laquelle nous avons choisi le corpus Subtitles en tant que source de données.

Le corpus Subtitles fait partie intégrale du corpus InterCorp qui est un ensemble des corpus parallèles dirigés par *Ústav Českého národního korpusu*.^{24, 25} L'intention du ČNK est de créer différents corpus notamment pour la langue tchèque, tels que le corpus SYN, qui est le corpus de référence du tchèque, les corpus diachroniques, et, parmi autres, le corpus parallèle InterCorp (voir ci-dessous). Ainsi, ČNK veut étudier la langue tchèque de nos jours sous différents points de vue (synchronique, diachronique, contrastive, dialectologique, « idiolectique », etc.). Le projet a été fondé en 1994 et, pendant à peu près ces dernières trente années, il a réussi à traiter plusieurs milliards (!) de mots textuels.

InterCorp est un projet sous gestion de ÚČNK, fondé en 2005 (cf. parmi d'autres Čermák & Rosen, 2012 ; Nádvorníková, 2017). Son but déclaré est de collectionner des textes authentiques écrits en tchèque et dans une autre langue – le plus souvent, on a affaire aux traductions. Il s'agit d'un corpus à versions (en tchèque « verzovaný korpus »), c'est-à-dire qu'ils existent plusieurs versions du même corpus dont la première (« version 0 ») a été rendue disponible en 2008 ; la plus actuelle, qui est la plus riche en fonctions et en données, date de janvier 2022 et dispose des textes en 41 langues. En 2022, l'ensemble d'InterCorp possède

²⁴ En français, *Ústav Českého národního korpusu* (ÚČNK) pourrait être traduit « L'Institut du Corpus National Tchèque ». Pour le corpus, nous employons l'abréviation ČNK (*Český národní korpus* ; « Le Corpus National Tchèque »).

²⁵ L'exposition suivante est basée sur le web de ÚČNK qui propose les informations actuelles sur les corpus en question. Le site est disponible sous <https://wiki.korpus.cz/doku.php/cnk:uvod> [consulté le 7 mars 2022].

1 572 000 000 mots. Les informations détaillées sur le(s) corpus sont à trouver soit en ligne sous <https://wiki.korpus.cz/doku.php/cnk:intercorp> [consulté le 7 mars 2022], soit dans Čermák & Rosen (2012).

Dans ce qui suit, nous ne traiterons que les éléments importants pour notre travail. Par rapport au présent mémoire (et à la didactique en général), InterCorp se fait remarquer notamment par les qualités suivantes :

- La version actuelle (14^e version, lancée en 2022) du corpus InterCorp contient 41 langues dont le français, l’anglais, l’italien, l’allemand, etc. Il s’agit donc d’une source extraordinaire non seulement pour les apprenants du français, mais pour les étudiants de toutes les langues communément enseignées dans les écoles tchèques/européennes.
- InterCorp convient notamment aux étudiants tchécophones, le tchèque étant la langue « pivot », c’est-à-dire que, dans ce corpus, toutes les occurrences dans n’importe quelle langue possèdent leurs équivalents en tchèque.
- Le niveau auquel les équivalences sont alignées est la phrase.²⁶ Cela nous paraît très pertinent pour l’emploi dans l’enseignement – le paragraphe étant une unité trop volumineuse, le mot au contraire trop « minuscule ». Cette observation concerne le but du mémoire, à savoir de développer la compétence communicative et de présenter le lexique uniquement dans les contextes authentiques.
- Autre qualité que nous croyons importante pour l’application du corpus dans le contexte scolaire, est la richesse du contenu de l’InterCorp. Parmi d’autres types textuels, le corpus contient p. ex. les textes juridiques, journalistiques, la littérature scientifique, les belles-lettres, etc. Donc, selon le but particulier, l’enseignant peut facilement trouver le type textuel qui lui convient.
- Ce qui peut se montrer très efficace en dressant les exercices, c’est l’annotation morphologique et syntaxique du corpus. Dans le présent travail, nous ne profiterons pas de cette qualité, mais nous la rappelons car elle peut aider p. ex. à créer les exercices sur la syntaxe.
- Comme nous l’avons déjà mentionné ci-dessus, il faut considérer également le côté « pratique » de l’application de corpus. InterCorp se fait remarquer

²⁶ Nous renonçons ici à la définition exacte (ou plutôt *bien réfléchie*...) de la phrase. Nous considérons la phrase dite « textuelle », donc n’importe quelle réalisation finie par une ponctuation finale.

par deux grands avantages par rapport à ce qui a été dit : c'est un corpus gratuit et très accessible du côté de l'utilisateur.

Dans le présent travail, nous ne nous servons que d'une partie (c'est-à-dire d'un subcorpus) du corpus InterCorp – à savoir, le corpus Subtitles. Il s'agit d'un corpus des sous-titres en plusieurs langues, dont le tchèque et le français. La taille de notre corpus touche presque 26 millions de mots.²⁷ La raison pour laquelle nous avons choisi ce corpus réside dans le but du présent travail – enseigner la langue parlée à l'aide de la langue authentique. En somme, le corpus Subtitles a ces avantages par rapport au but de notre mémoire :

- En tant que « sous-ensemble » du corpus InterCorp, il s'agit de corpus parallèle. Les occurrences y sont toujours disponibles en français aussi bien qu'en tchèque.
- Vu que les sous-titres accessibles proviennent de films (et non de documentaires, pièces de théâtre, etc.), il se montre adéquat de s'attendre à la langue parlée et actuelle – cette présupposition se montrera vrai dans la partie pratique.
- Le plus souvent, on a affaire aux sous-titres qui transcrivent les dialogues. Il n'est pas rare qu'une phrase dans le corpus réponde à une réplique dans le film. Ainsi, le matériel langagier s'approche du pôle de la langue conceptuelle parlée.²⁸ L'emploi didactique de la langue (quasi-)parlée a ses avantages évidents (exercices de « simulation », construction de dialogues, etc. – voir les chapitres ci-dessous).

4.2 Recherche et tri des occurrences

Comme nous l'avons déjà plusieurs fois mentionné, le corpus InterCorp n'est pas un corpus à but didactique. Pour cette raison, il faut, après la recherche dans le corpus, trier les occurrences et en choisir ceux qui sont pertinentes par rapport à notre objectif. Dans ce chapitre, nous décrirons une étape qui n'est pas typique pour le travail avec corpus, mais devrait l'être dans le cas d'exploitation de corpus pour les buts didactiques.

²⁷ Les informations détaillées sur le corpus sont disponibles sous <https://wiki.korpus.cz/doku.php/cnk:inter-corp:verzel3> ou sous <https://www.opensubtitles.org/fr> (le site de la base de données OpenSubtitles qui sert de source des données fournies par le corpus).

²⁸ La linguistique allemande, d'où cette classification provient, on parle de la *Mündlichkeit* – qui s'oppose à la *Schriftlichkeit*.

Nous avons commencé la recherche de données par l'enquête dans la langue CQL.²⁹ Étant donné que nous voulons travailler avec la construction verbo-nominale composée du verbe *avoir* suivi d'un substantif (sans déterminant, pour ne pas compliquer les choses...), il faut que le masque de recherche respecte les présuppositions mentionnées ci-dessous : dans la langue CQL, le verbe *avoir* en tant que lexème (c'est-à-dire sous toutes ses formes) se traduit par [lemma="avoir"]. Le fait que le mot suivant appartient à la classe des noms se marque par [tag="NOM"]. Et si l'on veut que les deux éléments suivent l'un l'autre, il ne faut que les juxtaposer, donc :

[lemma="avoir"][tag="NOM"]

Ensuite, nous avons cherché les occurrences correspondant à ce masque dans le corpus Araneum Francogallicum Maius³⁰ – pour vérifier les fréquences des collocations. Les résultats ont été rangés par leur fréquence absolue, afin de détecter ceux qui apparaissent le plus souvent. Il y avait p. ex. *avoir lieu*, *avoir besoin*, *avoir raison*, etc., mais aussi les occurrences bruit, p. ex. *avoir les*, *avoir @*, *avoir surement*, etc. La cause de ces bruits est, *grosso modo*, double : soit il s'agit de phrases « incomplètes » (telles que *avoir les*, qui manque un nom complétant le déterminant), soit les bruits sont causés par annotation défectueuse (p. ex. *avoir surement*, car le français ne connaît pas de nom « surement »).

Par conséquent, il a fallu trier les résultats « à la main ». Parmi les occurrences les plus fréquentes, nous avons indiqué ceux qui sont fréquentes et, en même temps, utilisable de point de vue didactique. En ayant ces deux critères devant les yeux, nous avons choisi les locutions suivantes (au nombre de 20) :³¹

1. *avoir accès*, B1
2. *avoir affaire* (mít s kým co do činění), B1
3. *avoir besoin* (potřebovat), A2
4. *avoir chaud*, A2

²⁹ « CQL » y est mis pour Corpus Query Language (angl. ; « langue de recherche dans le corpus »).

³⁰ Araneum n'a été employé que pour cette partie de la recherche – pour cette raison, le corpus n'a pas été présenté *supra*, n'étant pas source directe de données langagières. Il s'agit d'un corpus de web qui se fait remarquer par deux qualités : par sa taille énorme (à peu près 1,2 milliards de mots) et par la langue comprise – relativement proche de celle qui est l'objet du présent mémoire. Les informations détaillées à propos du corpus sont disponibles soit sous http://ucts.uniba.sk/aranea_about/index.html, soit dans l'article de V. Benko (2014).

³¹ Les indications de niveau de langue ne sont qu'approximatives – en général, elles respectent de nouveau les descripteurs du CCRL.

5. *avoir droit*³² (mít právo na), B1
6. *avoir envie* (mít chuť na, toužit po), A2
7. *avoir faim* (mít hlad), A1–A2
8. *avoir froid*, A2
9. *avoir hâte*, B1
10. *avoir honte* (styďet se), B1
11. *avoir intérêt*, B1
12. *avoir lieu* (konat se), A2–B1
13. *avoir mal* (bolet), A2–B1
14. *avoir peine*, B1
15. *avoir peur* (mít strach), A2–B1
16. *avoir raison* (mít pravdu), A2
17. *avoir recours*, B1
18. *avoir rendez-vous*, B1
19. *avoir soif* (mít žízeň), A1–A2
20. *avoir tort* (nemít pravdu, být na omylu, mýlit se), A2

C'est avec ce corpus de locutions que nous travaillerons par la suite.

Dans l'étape suivante, nous avons cherché ces locutions dans le corpus parallèle Subtitles. Le masque que nous avons employé a cherché deux lemmes (lexèmes) juxtaposés : *avoir* et le nom tiré de la liste ci-dessus. Dans la langue CQL, la requête p. ex. pour la locution *avoir intérêt* paraît comme suit :

[lemma="avoir"][lemma="intérêt"]

Vu que nous ne travaillons pas sur un corpus didactique, il fallait, dans cette étape aussi, trier les occurrences et trouver celles qui sont appropriées de point de vue didactique. Lors de ce tri, plusieurs facteurs ont décidé : longueur d'occurrence,³³ présence de traits de la langue parlée³⁴ et correspondance au niveau langagier A1–B1. Les occurrences choisies ont été téléchargées dans le tableur Excel, pour chaque locution à peu près 10 phrases parallèles.

Le tableur nous a servi en tant que « corpus didactique » avec lequel nous travaillerons dans la suite du mémoire. Les phrases particulières y ont été annotées pour qu'on puisse ensuite les classer. Les facteurs observés sont : le mot-clé représentant la locution

³² La locution similaire *avoir le droit de* citée dans les dictionnaires ne fait pas partie de notre corpus – sa forme ne correspond pas au masque CQL.

³³ Étant donné qu'il s'agit, *de facto*, du corpus de répliques (voir la description *supra*), ce critère était relativement facile à suivre.

³⁴ P. ex. absence des éléments forclusifs, questions *in situ*, etc. (voir plus haut).

concrète,³⁵ la difficulté, la complémentation, les traits de la langue parlée et une spécificité langagière potentielle. Dans les paragraphes suivants, nous illustrerons ces facteurs (excepté le premier) à l'aide d'exemples concrets.

La difficulté a été mesurée sur une échelle de trois degrés A, B et C – A étant les plus facile et C le plus difficile. Dans le tableau suivant, nous montrerons les phrases exemples qui appartiennent respectivement à toutes les classes nommées :

A : *J'ai chaud. – Je mi horko.* (phrase assez facile, correspond au niveau élémentaire)

B : *Ici, tout le monde a droit à son pourboire, c'est la loi ! – Spropitné dostávají všichni, ať si je zaslouží nebo ne. Je to zákon.* (phrase plus compliquée lexicalement (*pourboire, avoir droit*), mais grammaticalement relativement simple)

C : *On a eu chaud, mon vieux. – Páni, to bylo o chlup, kámo.* (sens idiomatique)

Les locutions du type *avoir* NP sont typiques par leurs complémentations variables. Dans le tableur, nous avons donc observé les prépositions régies : il y a les locutions régissant leur complémentation à l'aide de *de* (p. ex. *avoir besoin de*), *à* (p. ex. *avoir droit à*), *avec* (p. ex. *avoir rendez-vous avec*), etc., ou bien les locutions régissant un certain type de complément : CC de lieu (p. ex. *J'ai mal partout. – Všechno mě bolí.*), CC de temps (*La rencontre aura lieu ce vendredi 13. – Schůzka bude třináctého v pátek tento měsíc.*), etc., ou bien les locutions à « complémentation zéro » (p. ex. *Tu dois avoir faim. – Musíš mít hlad.*). Évidemment, la variation indiquée n'existe pas seulement *entre* les locutions, mais touche aussi les actualisations particulières de la même locution. À titre d'exemple, la locution *avoir tort* apparaît dans notre corpus didactique huit fois sans complémentation (p. ex. *Il a tort. – Plete se.*), mais trois occurrences régissent la préposition *de* (p. ex. *J'ai tort de faire ça... – Děláš chybu.*).

En ce qui concerne les traits de la langue parlée, nous avons marqué les occurrences dans lesquelles on a affaire « explicitement »³⁶ aux traits typiques du français parlé. Cette détection des occurrences typiques pour la langue parlée nous aidera à faire entrer celles-ci dans les exercices langagiers et, par conséquent, à familiariser les étudiants avec ce type

³⁵ Le mot-clé est le substantif suivant le verbe *avoir*, donc p. ex. pour la locution *avoir besoin*, le représentant est le mot « besoin ».

³⁶ Nos observations concernent notamment les traits mentionnés *supra* (chapitre 0), qui sont relativement faciles à relever et, en même temps, typiques pour la langue parlée (cf. aussi Blanche-Benveniste, 1997).

spécifique de la langue. Ainsi, en tant que « parlée », nous avons marqué p. ex. les occurrences suivantes : *T'as tort.* (élision faite sur *tu* ; cette « simplification » se montre très fréquente) ou bien *Tu as chaud, toi ?* (reprise du sujet postposée à la phrase sous sa forme tonique).

Finalement, ayant toujours devant les yeux l'objectif didactique du présent travail, nous avons indiqué toutes les occurrences qui peuvent, potentiellement, poser des problèmes aux apprenants non natifs du français. Étant donné notre contexte, nous nous sommes concentré sur les locuteurs-apprenants natifs du tchèque. Ce type d'observation ne concerne aucun plan de langue concret – il peut s'agir de difficultés :

- lexicales (p. ex. *J'ai rendez-vous demain, pour un boulot.* – bien que le mot *boulot* soit typique pour le français parlé, aucun des étudiants dans la classe où nous avons testé ces exercices ne le connaissait auparavant),
- phonétiques (p. ex. *J'ai besoin de riz, de haricots, de lait en poudre, de bière et votre meilleur whisky.* – ici, il faut, à ce que nous croyons, faire une remarque sur la phonétique du mot *whisky* [wiski], ou sur la prononciation « aspirée » de *haricots*),
- sémantiques (p. ex. *J'ai rendez-vous demain, pour un boulot.* – dans ce cas, il a fallu expliquer que le mot *rendez-vous* a une signification beaucoup plus large en français qu'en tchèque – le mot tchèque *rande*³⁷ ne s'emploie qu'au cas d'une relation amoureuse, tandis que son équivalent français peut désigner également un *rendez-vous pour un boulot*),
- ou bien pragmatiques (ici notamment en comparaison avec les phrases parallèles tchèques ; p. ex. *Ouais, tu as raison. Excuse-moi.* – *Jo, máte pravdu. Máte pravdu. Promiňte.* : en français, on tutoie, tandis qu'en tchèque, on vouvoie).³⁸

Un autre problème, que l'on rencontre forcément en tirant ses phrases exemple du corpus de la langue parlée, est la présence de gros mots. Les vulgarismes sont, dans ce type de documents, quasi omniprésents. On se pose la question s'il faut les « éliminer », ou, au

³⁷ Emprunté, d'ailleurs, au français.

³⁸ On ne s'intéresse pas ici au fait que dans cette réalisation concrète on emploie telle ou telle forme de politesse (vu que nous ne savons rien sur le contexte du film, des relations entre les personnages, nous ne savons même pas à partir de quelle langue les sous-titres ont-ils été traduits, ni quelle est la qualité de la traduction...). Mais, au contraire, nous pouvons, en tant qu'enseignants, très facilement profiter de cette non-correspondance en le tenant pour un point de départ pour les expositions conséquentes – dans ce cas-là, il s'agirait p. ex. du fait que le tutoiement et le vouvoiement sont utilisés dans les contextes légèrement différents dans les deux langues. Et cela est une partie légitime, voire nécessaire (cf. p. ex. RVP), de l'enseignement de n'importe quelle langue étrangère.

contraire, les traiter en tant que mots comme les autres. Cela nous semble être une question complexe : il faut prendre en considération l'âge des apprenants, le but de l'exercice concret, etc. Néanmoins, nous sommes partisan de l'approche qui a pour sa devise l'« intervention minimale ». Autrement dit, nous voulons présenter la langue « telle quelle » (tel est, d'ailleurs, le but du présent mémoire...) – il ne faut pas soigneusement cacher les vulgarismes, et, en même temps, il n'est pas nécessaire de les surestimer. Ce n'est que par hasard qu'aucun vulgarisme n'est parvenu à passer dans notre corpus didactique.

5 Types d'exercices

Dans ce chapitre pratique, nous présenterons les exercices tels que nous les avons testés au cours de notre pilotage (les détails seront présentés dans le chapitre 5.1). Comme nous l'avons déjà plusieurs fois mentionné dans les sections précédentes, l'objectif de nos exercices n'est pas seulement un phénomène (lexico-)grammatical concret (et plus en général la compétence langagière en tant que telle), mais aussi la capacité d'observer la langue authentique, et notamment les traductions « authentiques » où, relativement souvent, la version française ne correspond pas à la version tchèque. Pour les raisons articulées dans les parties suivantes, nous jugeons cette approche très utile de point de vue de l'école.

Pour ce qui concerne la méthodologie, nous avons exploité le corpus parallèle InterCorp v13, concrètement son « sous-corpus » appelé Subtitles (voir chapitre 4.1). Au moment où nous avons commencé à préparer les activités (en 2021), la version actuelle (v14) n'était pas encore disponible.

Le principe primaire d'organisation de ce chapitre est la division entre les activités *hands-off* (chapitre 5.1) et *hands-on* (chapitre 5.2). Les exercices *hands-off*, auxquels nous avons voué plus de place, se rangent en respectant le déroulement linéaire de l'enseignement : premièrement, nous présentons les activités d'introduction (ou « motivation », voir chapitre 5.1.1) ; deuxièmement, nous proposons plusieurs possibilités comment présenter les données de corpus aux apprenants (chapitre 5.1.2) ; ensuite on présente quelques exercices portant sur les capacités orales et écrites (les deux : productives et réceptives) ; pour « conclure » le processus d'enseignement, le chapitre 5.1.10 esquisse comment on peut préparer un test / une révision.

De point de vue de l'enseignant, il est très profitable de dresser (simultanément avec la préparation des exercices) la « version prof » qui contient un corrigé. Comment une version prof peut paraître est montré dans l'Appendix (Ex. prof). Étant donné qu'il s'agit d'un supplément facultatif par rapport au travail en classe, et que sa préparation suit celle présentée dans les chapitres suivants,³⁹ nous ne vouerons plus d'attention à cette problématique-là.

³⁹ En effet, de dresser une version prof est plus facile dans le sens qu'il suffit, p. ex. dans le cas des phrases à trous (voir chapitre 5.1.3), de « ne rien supprimer ». Dans le cas des dialogues ou d'écriture d'un texte (chapitres 5.1.4 et 5.1.8), la version prof n'a pas de sens, bien évidemment.

5.1 Exercices hands-off

Les exercices présentés dans cette partie reposent sur notre corpus didactique décrit dans le chapitre méthodologique. La plupart des exercices sont « à imprimer » ; néanmoins, la technique (notamment l'ordinateur et le projecteur) a sa place ici aussi.

Les activités ci-dessous ont été employées et testées à l'école (niveau lycée). Il s'agissait de différents niveaux langagiers (A0–B1), l'âge des apprenants était entre 15 et 19 ans.⁴⁰ Les remarques sur les activités concrètes seront présentées au cours du texte suivant.

5.1.1 Activité de « motivation »

Avant de commencer les exercices basés sur le corpus, qui portent sur un phénomène langagier concret, nous avons cru adéquat de préparer une activité « pour s'y plonger ». Sous l'influence de l'anglais (et vu que cette méthode s'applique notamment dans le milieu scolaire anglo-saxon... cf. *inter alia* Le Foll, 2021), nous avons appelé cette activité *motivation*.

L'objectif en est simple : il faut, avant de commencer p. ex. une explication de règle grammaticale, que les étudiants soient – ou *deviennent* – concentrés (« phase d'évocation »). Autrement dit, il nous faut une *captatio benevolentiae*. En effet, on peut commencer le cours par n'importe quelle activité qui, ce qui est important, se rattache au thème de la leçon en question, et laquelle introduction (motivation) nous amènera au sujet principal du cours.

Vu le sujet de nos exercices, nous avons commencé soit par une révision rapide de la conjugaison du verbe *avoir*, soit par une phrase contenant la construction *avoir* N (*Tout le monde a droit à son avis.*), de laquelle les étudiants ont été invités à discuter. C'est surtout la révision du verbe *avoir* qui s'est montrée très efficace : pour une fois, nous ne l'avons pas intégrée, ce qui a mené à une quantité *considérable* de fautes au cours des activités en classe. Mais une fois la révision était incluse, les étudiants ont fait remarquablement moins de fautes que le groupe où ce type d'activité a été omis. Également, aussi notre observation, il ne faut pas réviser uniquement les formes du présent, mais les formes les plus récurrentes (même dans les exempliers) – compris soit donc le passé composé, le conditionnel, etc. En parlant des formes (non seulement verbales), on fait face également à la prononciation. Aussi marginales et évitées qu'elles soient, il faut bien distinguer et correctement prononcer p. ex. les paires minimales comme *j'aurai* – *j'aurais* (prononciation fermée vs ouverte).

⁴⁰ Ce groupe d'étudiants correspond à ce qu'on appelle *vyšší gymnázium* (« lycée supérieur ») dans le système scolaire tchèque. Il s'agissait de différentes classes d'à peu près 5–15 personnes.

En général, l'emploi de ce type d'activité s'est montré d'être très efficace, notamment par rapport à la concentration des étudiants sur l'enseignement suivant et à l'évocation des informations déjà apprises, connues. Cette phase de motivation n'a, pourtant, son sens qu'au contexte de toute la leçon concrète – quelques minutes au début peuvent encourager les étudiants au travail dans le reste du cours.

5.1.2 Analyse de mini-corpus

Vu que nous travaillons avec un corpus en tant que source de données, il nous semble une conséquence naturelle que l'on présente une partie pertinente du corpus aux étudiants – avec l'objectif de les laisser l'analyser. Néanmoins, avant de le présenter, on doit inévitablement se poser la question suivante : Comment présenter les données ?

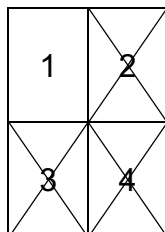
Il n'y a pas une seule réponse « correcte ». Dans cette partie, nous voulons proposer au moins quelques possibilités, qui ont été mises à l'épreuve dans l'école.

Les types de présentation peuvent être classifiés d'après plusieurs critères possibles ; ici, nous proposerons l'analyse selon deux critères suivants : la partie de corpus analysée par un étudiant et le médium de présentation. Toujours, il faut définir le corpus d'occurrences présentées aux étudiants – en le faisant, il ne faut pas oublier le but : soit nous voulons que chaque étudiant analyse toutes les occurrences (cf. Ex. 2), soit il est désirable qu'il ne travaille que sur une partie bien définie du corpus (cf. Ex. 1). Dans le premier cas, les étudiants peuvent travailler seuls ; dans le second, il est nécessaire que chacun présente les résultats de sa « recherche » aux autres, pour qu'ils puissent également apprendre la structure concrète.

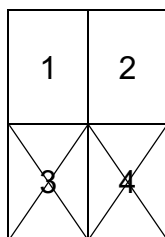
La classification d'après le médium de présentation offre, dans le cadre de nos exercices, trois modes possibles : les occurrences peuvent être présentées sous forme imprimée, dans une présentation PowerPoint, ou bien à l'aide des deux susdites possibilités en même temps. Pour les exercices en papier, il est possible de les envisager de plusieurs manières différentes – nous en avons testé deux : dans l'Ex. 1, nous avons présenté tout le corpus à la première page où chaque étudiant n'analyse qu'une partie du corpus. Le contraire est montré dans l'Ex. 2 : ici, la première page (tous les exercices sont censés à l'impression *recto verso*) se voit divisée en quatre parties, selon le schéma suivant :

1	2
3	4

Ensuite, il faut plier le papier horizontalement, puis verticalement – de manière que les étudiants, après les deux plis, ne voient que la partie 1 :

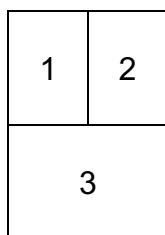


Dans le quadrant 1, on met les phrases en français avec lesquelles les étudiants travaillent dans la première étape, où ils essaient de les traduire et de deviner ce que les phrases veulent dire. Lors de la deuxième étape, ils peuvent vérifier leurs estimations en dépliant la feuille et en consultant le quadrant 2, sur lequel on a mis les occurrences équivalentes en tchèque. Ils disposent, au moment donné, de ces parties de la fiche :



L'ordre des occurrences dans la partie 2 est double : soit on peut les « aligner » aux phrases françaises, de telle manière que les apprenants voient tout d'abord les correspondances, soit on peut (comme dans l'Ex. 2) les présenter dans l'ordre aléatoire, donc non correspondant aux équivalents français – ici, le devoir est de chercher, en s'appuyant sur les estimations préalables, les correspondances.

Ensuite, les étudiants sont invités à déplier la feuille de nouveau et de consulter la partie 3, qui, pour sa part, est faite à partir de la fusion des quadrants 3 et 4 :



La partie 3 invite les étudiants à décrire et ranger les données (voir ci-dessous dans ce chapitre), et, en même temps, à les apprendre par cœur à l'aide des exercices (voir les chapitres suivants).

En ce qui concerne les diapositives PowerPoint, il est possible, comme nous l'avons déjà indiqué, de présenter seulement les occurrences sous cette forme (les exercices pratiques et la partie grammaticale ayant en papier) ou de se servir de diapositives pour toutes les étapes de l'apprentissage (présentation d'occurrences, analyse, exercices).

Après avoir présenté les occurrences parallèles, il faut que les étudiants les analysent. L'analyse, décrite de point de vue linguistique, se fait notamment aux niveaux suivants : on analyse, en général, la forme et le sens. Autrement dit, cette analyse basique correspond, *grosso modo*, à la théorie du signe langagier de F. de Saussure (1996), ou bien, plus récemment, à la grammaire de construction (« CxG ») qui, pour sa part, parle d'une *extended notion of the Saussurean sign* (voir Hoffmann & Dancygier, 2017). Il n'est pas étonnant que la CxG se voit parfois mise en œuvre même par les didacticiens de langues (cf. p. ex. De Knop & Gilquin, 2016) – et non seulement par ceux qui exploitent les corpus. Or, dans ce qui suit, nous nous servirons des exemples concrets liés à la construction *avoir N*.

Pour la forme, il faut que les étudiants apprennent à reconnaître les structures récurrentes, et, ce qui est encore plus important, qu'ils soient capables de généraliser. Dans nos exercices (cf. p. ex. Ex. 1), nous avons voulu qu'ils soient prêts à identifier la locution concrète avec sa complémentation prépositionnelle (c'est-à-dire la forme ; p. ex. *avoir besoin de*) et, pour le sens, la traduction en tchèque (p. ex. *potřebovat* pour l'exemple précédent). Schématiquement, cette relation apparaît comme suit :

FORME	français : <i>avoir besoin de QC</i>	tchèque : <i>potřebovat něco</i>
SENS	sens « partagé »	

Ensuite, il y a beaucoup de possibilités comment on peut traiter les résultats des analyses faites par les étudiants – toujours, ces activités doivent respecter la nature des occurrences : pour l'illustrer à un exemple concret, nous avons préparé les exercices portant sur les locutions *avoir N* qui régissent, parfois, un complément prépositionnel (p. ex. *avoir besoin de QC*). Vu que la plupart des locutions dans notre corpus régissent les prépositions *de*, *à* ou s'emploient sans complémentation (p. ex. *avoir raison*), l'exercice portant sur la classification

des locutions/occurrences⁴¹ d'après leurs compléments s'est montré très efficace pour la fixation et révision des prépositions appropriées. Cet exercice est incorporé dans l'Ex. 1.

Pour encore mieux illustrer comment les régularités de certains types d'occurrences peuvent être exploitées dans l'enseignement, nous faisons référence à la locution *avoir mal*. Vu la signification de la locution, il est évident que la plupart des occurrences seront complétées par la préposition *à* régissant un nom désignant une partie du corps humain (p. ex. *J'ai mal à la gorge.* ; *J'ai mal à la tête.* ; *J'ai mal au ventre.* – ces phrases sont prises du corpus didactique). Ainsi, à côté de l'apprentissage des locutions à *avoir* N, il est possible de réviser consciemment un domaine thématique du vocabulaire : les étudiants peuvent, p. ex., créer les phrases sur le modèle proposé par le corpus qui contiennent aussi les désignations d'autres parties du corps.

Après la phase d'exposition et d'analyse, il faut que les étudiants apprennent les locutions (ou, plus en général, les structures lexicales/grammaticales visées par les exercices), qu'ils les fixent. Pour une fixation efficace, nous nous sommes servi des phrases courtes, informatives (du point de vue lexico-grammaticale), fréquentes à l'usage, et faciles à retenir (donc p. ex. *J'ai faim.* – *Mám hlad.*). Si la locution concrète possède une complémentation prépositionnelle, il a fallu qu'elle soit incluse (p. ex. *J'ai envie de chocolat.* – *Dal bych si trochu čokolády.*). Après avoir fait les analyses et les exercices suivants, les étudiants ont été invités à apprendre ces phrases, soigneusement choisies, par cœur. Dans l'Ex. 1, les occurrences à mémoriser sont marquées en gras.

Néanmoins, il s'est avéré nécessaire de ne pas choisir uniquement les occurrences « faciles ». Au contraire, il a fallu ajouter quelques phrases plus compliquées pour qu'elles enrichissent les connaissances langagières des étudiants (lexique avancé, structures grammaticales plus complexes,⁴² etc.).

Ainsi, au cours du pilotage de nos exercices, les étudiants ont témoigné qu'ils connaissaient quelques locutions auparavant (p. ex. *avoir besoin*), mais que certaines leur étaient inconnues (p. ex. *avoir tort*). À côté de cela, ils ont bien apprécié les autres domaines « accompagnant » nos locutions : lexique, syntaxe, commentaires pragmatiques...

⁴¹ Dans certains cas, les prépositions ne varient seulement parmi les locutions, mais également entre les occurrences de la même locution. P. ex. *J'ai soif.* (sans complément) vs *J'ai soif de toi.* (complément introduit par la préposition *de*) – les deux exemples sont forgés.

⁴² Vu le but didactique des exercices, nous avons consciemment choisi les occurrences à traits de la langue parlée, laquelle caractéristique concerne tous les plans de langue, p. ex. la syntaxe : réductions (*T'as tort.*), questions *in situ* (*Tu as peur de quoi ?*), etc.

5.1.3 Phrases à trous

Après avoir présenté les données du corpus, nous passons vers les exercices dont l'objectif primaire est de fixer et réviser les connaissances acquises au cours de l'analyse. Le premier en sont les phrases à trous (voir les Ex. 1 et 2). Cet exercice, tel que présenté ici, n'a rien d'innovant. La seule spécificité, notamment par rapport aux textes/phrases à trous traditionnels, est qu'on ne travaille pas avec les phrases/textes monolingues, mais toujours avec les paires (*id est* avec les occurrences parallèles).

En conséquence, on a plus de « possibilités » où les mots peuvent être supprimés. Prenons, en tant qu'exemple, la « bi-phrase » suivante : *Tu as honte de travailler avec nous ? – Stydiš se pracovat s námi?* Ici, on peut supprimer un/des mot(s) en tchèque, en français, ou bien dans les deux langues :

Tu as honte de travailler avec nous ? – Stydiš se _____ s námi?

Tu as _____ de travailler avec nous ? – Stydiš se pracovat s námi?

Tu as _____ de travailler avec nous ? – Stydiš se _____ s námi?

L'idée principale de cet exercice est que les étudiants remplissent les lacunes (« trous ») en consultant les traductions. Bien que ce dialogue entre les langues se soit avéré très fructueux, il posait parfois des problèmes : les étudiants n'étaient pas accoutumés de travailler avec ce type d'occurrences, donc ce qui s'est montré la difficulté principale était la non-correspondance entre les phrases (p. ex. *On a accès à la boîte mail de Ross. – Nabořili jsme se Rossovi do mailu.*). D'après les résultats de notre pilotage, cela pose problèmes tout au début d'exploitation de corpus – après quelques reprises, le problème disparaît progressivement : les apprenants se sont habitués à ce type d'exercices, et notamment aux occurrences provenant du corpus et ses spécificités. Dans une enquête informelle à la fin du pilotage, les étudiants ont apprécié ce fait de non-correspondance des occurrences qui les « faisait réfléchir ».

À l'instar des moules présentés ci-dessus, il est possible de créer les exercices portant sur les phénomènes grammaticaux ou lexicaux. Ce qui s'offre par rapport à notre projet c'est, par exemple, une activité visant les rections prépositionnelles, qui sont récurrentes dans notre cas :

Tu as honte ____ travailler avec nous ?⁴³

Vous ignorez ____ qui vous avez affaire.

⁴³ Ici, il n'est pas nécessaire d'ajouter les traductions dans l'autre langue.

5.1.4 Dialogues

Les données prises du corpus peuvent être exploitées de diverses façons, en tant que point de départ pour d'autres activités. Dans le présent mémoire, nous exploitons le corpus de sous-titres – vu leur nature textuelle (dans la quasi-totalité des cas, les sous-titres proviennent de dialogues), il s'avère possible d'en profiter et de créer des exercices basés sur cette caractéristique.

Dans les Ex. 1 et 2, nous en proposons la forme concrète. Dans notre corpus didactique, nous avons marqué les occurrences à partir desquelles on peut facilement construire un dialogue – nous avons donc choisi surtout les questions (p. ex. *J'ai raison ?*), les phrases exclamatives (p. ex. *J'ai mal à la lèvre !*), ou bien les phrases employables dans la vie quotidienne (p. ex. *Moi aussi j'ai envie d'aller à la plage.*). Le seul critère, nous le répétons, était le « potentiel conversationnel » de ces phrases.

Pour un exercice, nous avons choisi quelques 5–7 phrases.⁴⁴ Dans la classe, les étudiants ont été invités à lire les phrases, à les comprendre (ou traduire, si nécessaire...) et à créer à deux ou dans un petit groupe un dialogue construit sur les phrases proposées.

5.1.5 Discussion

Une autre activité visant la production orale repose sur la discussion en classe. Ici, nous ne traiterons pas cette activité très en détail – sa liaison aux exercices exploitant le corpus est relativement libre. Néanmoins, une discussion s'est avérée un bon complément des activités présentées dans ce chapitre-là.

Nous avons conçu cet exercice comme une discussion autour d'une thèse principale (voir Ex. 2). Il faut trouver une phrase ou un texte court qui touche le thème de la leçon concrète – peu importe si le thème est d'ordre « thématique » (p. ex. boissons, animaux), ou bien d'ordre grammatical (p. ex. plus-que-parfait). Dans notre cas, cette cohérence s'est faite au niveau lexico-grammatical – l'Ex. 2 offre la phrase suivante en tant que point de départ de discussion : *Tout le monde a droit à son avis.*

Quoique cette phrase soit tirée du corpus, le critère décisif ici n'est pas la source de données (il n'est pas toujours possible de trouver une phrase pertinente dans le corpus...), mais le potentiel de la phrase d'enflammer une discussion. Donc, il devrait s'agir d'une phrase « frappante », concernant, le cas idéal, soit la situation actuelle, soit un problème, un dilemme

⁴⁴ Sans traduction tchèque, vu qu'elles sont faciles du point de vue langagier ; néanmoins, si nécessaire, il est possible de les ajouter, bien sûr. Cela notamment dans les cas où l'emploi de la phrase risque de mener à une confusion, voire malentendu (particulièrement de sorte pragmatique) ; p. ex. *On a eu chaud, mon vieux.* – *Páni, to bylo o chlup, kámo.*

perpétuel et/ou philosophique. D'ailleurs, cette activité peut profitablement être employée comme motivation (cf. le chapitre 5.1.1).

5.1.6 Repérage des traits du français parlé

Pendant l'analyse du mini-corpus, présentée dans le chapitre 5.1.2, il est possible de prévenir les étudiants des phénomènes « marginaux » par rapport au thème principal des exercices, la construction *avoir* N. Vu qu'il s'agit d'une activité à part, qui peut facilement devenir autonome, nous la traitons séparément dans ce sous-chapitre.

Le but de notre projet est de proposer les exercices sur les locutions *avoir* N, en les présentant dans le contexte de la langue parlée. Les locutions *avoir* N sont notre objectif primaire, le français parlé étant plutôt un but implicite. Néanmoins, on peut l'explicitier : dans l'Ex. 3, nous avons encouragé les étudiants à indiquer les occurrences contenant n'importe quel trait de la langue parlée. Dans la plupart des cas, les étudiants pouvaient sans effort trouver les traits parlés ; sinon, la traduction tchèque fournie par le corpus était secourable. Il s'agissait avant tout de phénomènes grammaticaux (p. ex. absence des éléments forclusifs, questions *in situ*, etc.).

5.1.7 Traduction

À partir du corpus parallèle, on peut très facilement créer les exercices de traduction – en supprimant la partie équivalente. Ainsi, la traduction en classe peut être bidirectionnelle, ce qui est montré dans l'Ex. 1.

Parfois, la traduction est jugée « démodée » par certaines approches didactiques (cf. p. ex. l'approche communicative) ; néanmoins, elle s'est avérée efficace lors de notre pilotage : parce que les données du corpus ne sont pas « forgées » (c'est-à-dire préparées avec l'objectif didactique), elles sont très concrètes, tirées de la communication réelle. Pour cette raison, il nous semble peu probable que les étudiants trouvent la traduction disponible dans le corpus (p. ex. *J'ai envie de me promener. – Asi se pūjdu projít.*) – cette présupposition s'avérerait valide au cours du pilotage. Bien sûr, il est possible de « guider » les apprenants comment ils doivent traduire les phrase concrètes : dans notre cas, la seule condition/restriction de la traduction était la nécessité d'employer toujours la locution *avoir* N (pour la traduction vers le français) ; la traduction français–tchèque restait sans restriction.

L'idée générale de cette activité est de développer la créativité des étudiants et leur faculté de résoudre vite diverses situations (quasi-)communicationnelles.⁴⁵ Pour cela, nous

⁴⁵ Puisque nos données proviennent du corpus, elles ne sont pas toujours très clairement contextualisées, et leur sens « communicationnel » ne doit pas être toujours évident.

n'avons vu les phrases équivalentes du corpus que comme un « guide », une source d'inspiration – la créativité des étudiants, qu'ils ont montrée en traduisant, était appréciée.

5.1.8 Écrire un texte

Une autre compétence que l'on peut entraîner à l'aide des données du corpus est la production écrite. En général, il est possible de procéder de manière similaire comme dans le cas de l'entraînement de la production orale (cf. chapitre 5.1.4). On présente aux étudiants un corpus de phrases à partir desquelles ils sont censés de rédiger p. ex. une histoire écrite. Cette procédure a été employée dans l'Ex. 2.

Sinon, pour développer autrement la langue écrite, on peut proposer aux apprenants seulement les parties de phrases à partir desquelles ils écrivent un texte – p. ex. *Lucien avait envie de pleurer parce que...* (l'occurrence étant modifiée). Une autre possibilité est de compléter les phrases (ainsi, on ne vise pas le niveau textuel, mais « seulement » le niveau de la phrase) – p. ex. *Lucien avait envie de...*

Évidemment, cet exercice ne présente que le pôle opposé de l'activité « phrases à trous » (voir le chapitre 5.1.3). Lors du pilotage, nous avons testé plusieurs degrés du continuum « rempli–vide » – tous se sont avérés appropriés, bien que pour les buts différents.

5.1.9 Révision d'un phénomène grammatical

Dans le chapitre 5.1.2, nous avons déjà fait allusion à la récurrence de certains phénomènes linguistiques. La capacité de trouver les relations syntagmatiques aussi facilement que paradigmatisques (cf. p. ex. Čermák, 2013 ; Čermák, 2017a) est un des atouts indéniables du corpus – par conséquent, le corpus devient une source idéale pour les exercices portant sur les récurrences. Ce sous-chapitre sera consacré d'une part aux phénomènes récurrents (parfois, on parle de la collocabilité ; une explication détaillée p. ex. dans Čermák, 2017a), de l'autre aux phénomènes non récurrents (« difficultés »).

Pour ce qui concerne la récurrence dans le corpus,⁴⁶ elle est de double sorte : lexicale et grammaticale. La récurrence lexicale a déjà été présentée dans le chapitre 5.1.2 sur l'exemple de la collocabilité de la locution *avoir mal* – les actualisations ne sont attestées qu'avec un complément désignant une partie du corps.

Le même pour les récurrences grammaticales : les locutions du type *avoir N* sont typiques pour un régime prépositionnel, les plus fréquentes étant les prépositions *de* et *à*. Dans le corpus, ces deux prépositions (ou plutôt leurs phrases prépositionnelles) occur-

⁴⁶ La généralisation proposée repose sur les observations de notre corpus et, en même temps, elle est influencée de la distinction traditionnelle de la didactique de langue étrangère en trois domaines : grammaire, lexique et phonétique (Choděra, 2013).

relativement souvent, sous forme de pronoms adverbiaux (c'est-à-dire *en* et *y*). Un exercice visant ce phénomène grammatical fait partie de l'Ex. 1 (révision de COD et COI). L'avantage de ce procédé est double : d'une part, c'est l'entraînement systématique, de l'autre le fait que les apprenants voient et apprennent comment on emploie le phénomène en contexte.

La révision systématique (« récurrente ») de la phonétique se trouve à part, vu la diversité formelle des collocants. Néanmoins, dans le cas des locutions *avoir* N, il a fallu – à tous les niveaux de connaissances – réviser la prononciation des formes du verbe *avoir* : par exemple, un problème récurrent des apprenants tchèques est la distinction correcte entre les désinences ouvertes et fermées, *j'aurai* [-e] vs *j'aurais* [-ε].⁴⁷

À côté des récurrences, il est utile de faire attention aux traits problématiques et/ou difficiles. Il faut penser à cela déjà en préparant le corpus didactique. Nous avons, en triant les données, marqué les occurrences que nous avons jugées difficiles – peu importe, s'il s'agit de difficultés lexicales, grammaticales, phonétiques ou bien pragmatiques (voir le chapitre 4.2, où le procédé est esquissé plus en détail). Bien sûr, il faut que l'enseignant soit toujours prêt à expliquer à l'improviste les difficultés, mais il est avantageux de les anticiper, au moins partiellement.

5.1.10 Test

Après avoir exposé le thème en question, il est possible de tester (soit pour réviser, soit pour évaluer) les connaissances acquises.⁴⁸ Le test, lui aussi, peut être dressé à partir du corpus. Les données de départ, aussi bien que la méthodologie et typologie d'exercices peuvent être reprises des chapitres précédents – phrases à trous, traduction, etc. L'avantage didactique du test basé sur le corpus est sa complexité (voir ci-dessous) et, en même temps, certain degré de « prédictabilité » pour les étudiants qui, après avoir absolu l'exposition et avoir fait les exercices, connaissent bien la forme.

Toujours, avec la méthodologie prises dans ce projet, le test portera, comme tous les exercices, sur un phénomène langagier concret. Néanmoins, ils existent deux manières de procéder en préparant le test. Soit on teste uniquement la connaissance du phénomène, soit on teste la connaissance du phénomène ancré dans un contexte plus large (ce que nous appelons la « complexité » du test).

⁴⁷ L'explication du fait n'est pas compliquée : le tchèque ne dispose pas de cette distinction. Pour plus d'exemples, cf. Dohalská & Schulzová (2015), Dohalská & Suková Vychopňová (2015) ou Fenclová (2010).

⁴⁸ Ce test est adéquat notamment si l'on cible la connaissance d'un phénomène concret de la langue.

Pour le premier cas, l'exercice avec les phrases à trous est une solution possible (voir l'Ex. 4-A). P. ex., on peut supprimer dans chaque occurrence de la locution *avoir N* le nom, en offrant les traductions pour que les étudiants sont capables de reconstruire la phrase :

J'ai _____ à une explication. – Dlužíte mi vysvětlení.

L'autre possibilité est le test complexe, c'est-à-dire un test qui ne teste pas seulement le phénomène concret (dans notre cas les locutions *avoir N*), mais aussi la capacité de pouvoir correctement l'employer dans son contexte « naturel ». Pour ce type de test, nous nous sommes servi de deux exercices : traduction bidirectionnelle et la recherche des équivalences (voir l'Ex. 4, parties B et C).

Le test présenté dans l'Appendix (Ex. 4) a également été employé lors du pilotage, dans la classe où nous avons exposé les locutions à l'aide de l'Ex. 1. Les résultats de cette expérience⁴⁹ étaient satisfaisants : les étudiants ont obtenu entre 16 et 20 points, la plupart d'eux 19 ou 20.

Pour tester la capacité de produire les tests (soit à l'écrit, soit à l'oral), il est possible de demander aux étudiants de préparer une simulation (en employant certain nombre de locutions), de rédiger un texte, etc.

En analysant les résultats du test, nous avons demandé aux étudiants leurs expériences avec la révision pour le test. La plupart d'eux ont vivement apprécié les phrases exemples qu'ils ont apprises par cœur⁵⁰ (voir le chapitre 5.1.2).

5.1.11 Résumé

Les exercices présentés dans ce chapitre ont visé trois des quatre activités langagières : l'écriture (c'est-à-dire la production écrite), la lecture (la compréhension écrite) et l'oral (la production orale). Dans le tableau suivant, nous résumerons les activités ; leur développement possible est toujours illustré par une activité exemple mentionnée ci-dessus :

	Production	Réception
Orale	Dialogues	?
Écrite	Rédaction de texte	Analyse de mini-corpus

⁴⁹ Le test a eu lieu le 15 février 2022 au lycée Gymnázium Jana Keplera (à Prague). Le niveau des étudiants était A1.

⁵⁰ En général, il se recommande d'apprendre ces locutions sous forme de *chunks* – cette forme d'apprentissage s'est avérée déjà plusieurs fois bénéfique, cf. p. ex. Choděra (2013).

Il se pose la question sur la nature de la réception écrite – qu’est-ce qui pourrait remplacer le point d’interrogation ? Évidemment, il y a plusieurs réponses possibles ; mais par rapport au corpus en tant que source de données pour l’enseignement/apprentissage du français, il serait opportun à penser aux corpus de la langue parlée. Pour cette problématique, nous faisons référence aux études correspondantes, p. ex. Detey et al. (2011) ou Le Cunff (2005). De plus, les activités présentées peuvent facilement couvrir toutes les étapes de l’enseignement (selon Chodéra, 2013) : l’exposition, la fixation, l’application et le contrôle.

Néanmoins, les données de corpus ne sont pas omnipotentes : il faut, au cours de l’enseignement, mentionner les difficultés, ou bien corriger les fautes les plus graves que les étudiants font. Par exemple, lors du pilotage, il a fallu expliquer les différents emplois de la locution avec le mot *chaud*. Nous avons procédé à partir de la « paire minimale » *j’ai chaud* et *il fait chaud*, que les apprenants confondaient assez souvent. Ce type d’information n’est pas, dans notre « mode » de travail, à trouver dans le corpus – l’enseignant est censé de profiter de sa créativité et, avant tout, de son expérience.

Également, il s’est avéré utile de prendre en considération la collocabilité du phénomène concret et de confronter les deux langues (maternelle et étrangère), si possible. Nous ne l’avons pas intégré dans nos exercices, mais il est avantageux, voire nécessaire, d’avertir les étudiants aux différences entre le français et l’anglais – vu le fait que les connaissances de l’anglais sont souvent beaucoup plus élevées que celles du français. Bien qu’inopinément, nous étions, au cours du pilotage, forcé à faire face aux confusions de ces deux langues étrangères ; surtout, il s’agissait des confusions lexicales (dues à l’homonymie étymologique anglo-française), telles que p. ex. *formidable* qui, en anglais, possède une signification antonymique....

Pour être objectif, il faut également rappeler les inconvénients des exercices présentés dans ce chapitre. En général, ils correspondent aux désavantages discutés dans la partie théorique (voir le chapitre 2.2) – nous ne les mentionnerons que brièvement : chronophagie (concerne notamment la préparation), nouveauté pour les étudiants (il faut que les apprenants s’habituent à ce type particulier d’exercices),⁵¹ ou bien la nécessité d’assistance pédagogique tout au début.

⁵¹ C’est notamment la non correspondance littérale des concordances parallèles.

5.2 Exercices hands-on

Dans cette partie, nous nous occuperons un peu des exercices *hands-on*, c'est-à-dire des exercices pendant lesquels les étudiants travaillent directement avec le corpus. Il est évident que cette approche peut avoir plusieurs avantages ainsi que certains inconvénients (nous les avons vus et discutés dans les chapitres 2.1 et 2.2).

Pour l'approche *hands-on*, il semble nécessaire de familiariser les apprenants avec la recherche dans le corpus. Comme nous l'avons dit déjà dans les chapitres précédents, la recherche dans les corpus sous la gérance de *Ústav Českého národního korpusu* ne diffère pas remarquablement de celle menées à l'aide des moteurs de recherche comme p. ex. Google.⁵² Il est hautement probable, de nos jours, que chaque apprenant sait s'en servir.

Ce qu'il faut apprendre aux étudiants avant tout c'est le travail basique avec le corpus. Notamment les activités suivantes sont à souligner avant de laisser les apprenants se plonger dans le corpus :

1. Indiquer l'emplacement du corpus (dans notre cas, c'est le site officiel du ČNK, <https://korpus.cz>) et expliquer (en le montrant en même temps !) comment les étudiants peuvent accéder au corpus concret (ici : InterCorp). Nous pouvons régler cette indication par une simple navigation arborescente, donc p. ex. « Page d'accueil / KonText / corpus : InterCorp v14 – Czech ou French ».⁵³
2. Instruire les étudiants quels paramètres il faut respecter. Pour nous, c'est uniquement la nécessité d'avoir choisi le corpus parallèle franco-tchèque (ou *vice versa*).
3. Expliquer les notions de l'interprétation d'une « requête simple » : avant tout, il faut montrer comment le concordancier interprète les mots cherchés. P. ex. si l'on cherche le mot *pes* (en tchèque « chien », forme du lemme), le concordancier répond par toutes les occurrences (fléchies) du mot *pes*, donc aussi *psi*, *psům*, *psovi*... Au contraire, si l'on écrit *psům*, le concordancier ne cherche/trouve que les occurrences du datif pluriel.

⁵² Cette assertion ne concerne que la recherche « de base ». Le grand avantage de ces corpus est d'un côté l'accessibilité considérable pour les débutants, de l'autre la possibilité d'effectuer les recherches les plus compliquées et complexes. Mais ce second point n'a pas d'importance pour nous, dans ce cas-là.

⁵³ Nous citons la version actuelle, donc la version 14 (2022).

4. Finalement, on devrait montrer aux étudiants comment interpréter les résultats trouvés – afin de les protéger de la confusion dans les données (cf. chapitre 2.2).

5.2.1 Phraséologie français–tchèque

Le premier exercice *hand-on* que nous présenterons vise avant tout les niveaux plus avancés, étant donné que le contenu en sont les phraséologismes et diverses locutions figées. En tant qu'exemple nous utiliserons une locution figée qui fait partie de notre corpus : *avoir chaud* au sens non compositionnel.

On peut laisser les étudiants chercher, dans ce cas-là, une réalisation concrète qui, dans nos recherches préparatoires, s'est montrée relativement fréquente, p. ex. la phrase *On a eu chaud*. Le but est simple ici : les apprenants doivent parcourir les textes parallèles et chercher la signification de la phrase contextualisée.

Le corpus offre, parmi autres, les équivalents suivants :

To bylo těsně.

To bylo o fous.

Ted' jsme měli namále.

Bylo to jen tak tak.

Na poslední chvíli. etc.

Ce que ces traductions montrent est la quantité énorme des équivalences possibles. Les avantages sont évidents : l'apprenant voit la construction dans plusieurs contextes et il rencontre différentes possibilités de traduction, ce qui augmente la vraisemblance du stockage dans la mémoire.

Hors du cadre du mémoire, nous montrerons ce type d'exercice sur le phraséologisme très souvent cité, *donner sa langue au chat*. Si nous le tapons littéralement dans le concordancier, la réponse n'est pas si « vaste » (vu la spécificité relative de l'enquête – p. ex. *sa* est un pronom pour une personne concrète, *donner* était placé à une position exacte par rapport à la suite de la phrase, et ainsi de suite). C'est pourquoi il faut élargir (« déconcrétiser ») le champ de l'enquête : ici, il est possible de le faire en omettant le verbe et le pronom possessif, d'où on obtient l'enquête *langue au chat*. Dans ce cas-là, les résultats sont beaucoup plus prometteurs (et nombreux, soit dit en passant...) :

Vzdávám se.

Dáš se poddat?

Nemám vubec páru., etc.

Pour ce qui concerne la morphosyntaxe de l'idiome, le corpus nous montre (grâce à l'omission des deux composantes) que la forme indiquée *supra* n'est pas la seule possible. Dans notre corpus, elle n'est même pas la plus fréquente. Ainsi, on avait affaire à :

Je donne ma langue au chat.

Ma langue au chat !

Langue au chat !, etc.

Parmi quelques occurrences bruit, on peut également trouver des occurrences qui attestent bien l'actualisation du phraséologisme, lesquelles occurrences sont, d'ailleurs, intéressantes du point de vue didactique. Ainsi p. ex. la phrase suivante :

Martin avoue qu'il donne sa langue au chat. – Martin přiznává své rozpaky.

L'exemple ci-dessus nous apprend au moins comment un phraséologisme concret peut être syntaxiquement intégré dans une phrase plus complexe,⁵⁴ et comment on peut le traduire en tchèque⁵⁵ sans se servir de la traduction courante *dát se poddat*.

5.2.2 Phraséologie tchèque–français

Cette partie présente, en effet, la même chose que le chapitre précédant : la recherche de la phraséologie, cette fois-ci dans la direction opposée, tchèque–français. La raison pour laquelle nous avons ajouté ce petit chapitre est le fait que la direction de recherche « compte ». Dans le chapitre 5.2.1, nous avons effectué la recherche dans la direction français–tchèque – en supposant que le français soit la langue à apprendre, c'est-à-dire non maternelle. Le tchèque, au contraire, se voit supposé d'être la langue maternelle : pour cela, nous n'avons pas commenté la classification et l'évaluation de données obtenues, car il n'y a pas de risque de faute langagière.

Si l'on cherche, pourtant, les équivalents dans une langue non maternelle, il faut être beaucoup plus attentif – l'apprenant est censé ne pas connaître la langue parfaitement, donc il se voit exposé au risque de fautes sur tous les plans langagiers. La linguistique de corpus a,

⁵⁴ Dans ce cas en fonction d'une complétive.

⁵⁵ Ici, nous laissons de côté la direction de la traduction, sans parler des « tierces langues ».

bien évidemment, ses moyens pour prévenir ces fautes. Néanmoins, dans le contexte scolaire, nous ne pouvons qu'employer les règles les plus élémentaires : par exemple, on peut dire aux étudiants tout simplement qu' « un hapax ne compte pas ». Autrement dit, pour être sûr que l'on a affaire à une construction « neutre », on a besoin d'un ensemble d'attestations le plus vaste possible.

Pour concrétiser les idées présentées jusqu'ici, nous proposons un exercice fondé sur la recherche des équivalents de l'idiome tchèque *bylo to o chlup*. De nouveau, si l'on cherche la phrase *bylo to o chlup*, on n'obtient pas de résultats. Pour cette raison, il ne faut que chercher la locution *o chlup*. Le concordancier répond par 63 occurrences.

En général, nous croyons qu'il faut profiter du corpus parallèle – c'est pourquoi nous proposons de diviser cet exercice en deux parties : la partie portant sur le tchèque, et celle qui porte sur le français. Bien que nous nous trouvions dans le cadre de l'éducation de langue étrangère, il faut pourtant toujours respecter la langue maternelle. Et ce qui est de plus : l'employer en tant que moyen d'entraînement du savoir-faire analytique.

En parlant des occurrences tchèques, il serait utile si les apprenants savent analyser et décrire « ce qui se passe », sans se servant de la terminologie linguistique, bien sûr. Par exemple, les étudiants peuvent constater que la locution *bylo to o chlup* n'a point besoin du verbe copule *být* (« être ») – en plus, dans la plupart des cas, il ne s'agit pas de prédicatif, mais d'un complément circonstanciel : p. ex. *K tomu ráno o chlup minul dvě myši a měl pocuchané sebevědomí.* (*Ce matin-là, en outre, il avait raté coup sur coup deux souris, en souffrait dans son orgueil.*).

Pour ce qui concerne le corpus des phrases françaises, les étudiants cherchent, tout d'abord, les équivalents. À notre avis, il faut indispensablement encourager les apprenants à être prudents face aux équivalents potentiels – comme nous l'avons déjà proposé, la fréquence absolue (sous le contrôle continu de l'enseignant !) semble d'être critère suffisant par rapport aux buts suivis.

Les équivalents récurrents, dans le corpus de nos 63 occurrences, étaient p. ex. :

- *de justesse* (*De justesse ! – To bylo o chlup.*)
- *l'échapper belle* (*Cette fois, je crois qu'on l'a échappé belle. – Myslím, že tentokrát jsme vyvázli o chlup.*)
- *faillir* (*Ca [sic] a failli mal tourner. – To bylo o chlup.*)

Il devient clair que ce type d'entraînement langagier peut bien enrichir le vocabulaire d'étudiants. Néanmoins, ce type d'exercice n'est pas « universel », il ne convient pas aux tous les niveaux. De plus, il faut que les apprenants soient sûrs des bases de la langue, parce que le matériel (ce qui est encore plus le cas en travaillant avec le corpus Subtitles...) n'est pas toujours « parfait », laquelle remarque concerne notamment l'orthographe (voir aussi l'exemple ci-dessus).

5.2.3 Résumé

Étant donné que ce mémoire vise surtout les exercices *hands-off*, nous n'avons présenté le type *hands-on* sur les corpus parallèles que très brièvement – dans ce domaine, il reste évidemment beaucoup de place pour de nouvelles idées et innovations. Ce que nous avons observé, tout « en passant »,⁵⁶ sont les faits suivants :

- Ce type d'activité est très chronophage, tout à l'instar de la littérature citée dans le chapitre 2.2. En comparaison avec le type *hands-off*, sa chronophagie est de nature différente : tandis que dans le cas des exercices *hands-off* elle concerne le temps nécessaire pour la préparation, dans le cas du type *hands-on* elle touche le travail en classe.
- Par conséquent, ces activités ne sont pas « employables » lors de chaque cours.
- Cette activité nécessite la présence permanente de l'enseignant – il faut que le maniement de données soit *toujours* contrôlé (risque de « désorientation » dans le corpus, problèmes techniques, etc.).
- Les exercices *hands-on*, en général, sont plus utiles pour les apprenants avancés – d'après notre expérience, notamment pour les étudiants universitaires (en linguistique, le cas idéal...). Pourtant, même au niveau secondaire, on peut bien les utiliser (pour enrichir le vocabulaire, pour thématiser les structures grammaticales moins courantes, etc.). L'obstacle le plus évident est, à notre avis, de nouveau le fait que les étudiants ne sont pas habitués à travailler de cette manière-là.

À l'instar des études effectuées (cf. André, 2020), nous faisons attention à la possibilité d'utiliser les corpus pour consulter les doutes sur la langue étrangère, ou même sur sa propre langue maternelle. Cette information peut être utile p. ex. en écrivant une dissertation

⁵⁶ Les exercices *hands-on* n'étaient pas, contrairement aux exercices *hands-off*, systématiquement testés à l'école.

où, notamment dans le cas d'une langue étrangère, on a souvent besoin de vérifier divers phénomènes de langue.

Pour conclure, nous croyons très utile de travailler, de temps en temps, avec le corpus de la manière *hands-on*. Les arguments généraux parlent du développement de la « conscience linguistique » non seulement dans la langue étrangère, mais aussi dans la langue maternelle, du développement de la pensée analytique, de la faculté de travailler avec des données, etc.

6 Conclusion

Pour conclure le mémoire, revenons à son début où, dans le chapitre d'introduction, nous avons jalonné le chemin que nous avons voulu prendre : on a décidé de contribuer au domaine de la didactique au niveau secondaire basée sur le corpus parallèle franco-tchèque. Au cours du texte, nous avons proposé, en s'appuyant sur les études du domaine déjà effectuées, quelques activités et procédés possibles à prendre lors de l'enseignement scolaire.

Dans le premier chapitre théorique (chapitre 2), nous avons présenté le corpus en tant qu'ensemble numérisé de textes et traité son potentiel dans la didactique de « deuxième langue étrangère » – cela notamment par rapport aux exigences du système scolaire tchèque. Au regard du RVP, le corpus s'est avéré (théoriquement et pratiquement) un outil relativement puissant, dont on peut, sans aucun doute, bien profiter au cours de l'enseignement. Du point de vue du RVP, le corpus recouvre plusieurs compétences clé, ce qui est plus en détail décrit dans le chapitre 2.3.

Avant de rappeler les avantages et inconvénients du corpus, il ne faut pas oublier de parler de la motivation. Ici, la motivation est double : la motivation d'apprenants et la motivation d'enseignants. Ce qui s'est montré de prime importance est la seconde mentionnée – nous tirons cette conclusion non seulement de la littérature scientifique, mais aussi de nos observations dans l'école. Donc nous croyons que c'est ici que l'on devrait commencer à « propager » les corpus.

Pour ce qui concerne les avantages de l'exploitation didactique de corpus, dans le chapitre 2.1, nous avons cité p. ex. la motivation d'apprenants, le développement des capacités analytiques, de la pensée critique, etc. Notre pilotage a montré que les étudiants appréciaient notamment le fait qu'il s'agissait de la langue authentique et qu'ils ont appris un phénomène de langue à l'aide des occurrences récurrentes ; autrement dit, qu'ils l'ont appris dans plusieurs contextes.

Les désavantages, cités dans le chapitre 2.2, étaient aussi présents : du point de vue de l'enseignant, l'inconvénient le plus grave est la chronophagie – il faut relativement beaucoup de temps pour pouvoir préparer les activités. Pour les apprenants, le plus grand obstacle était le fait qu'ils ne sont pas habitués à ce type de travail. Les deux désavantages se sont avérés cardinaux dans le pilotage.

Afin de pouvoir tester l'application de corpus à l'école, il est nécessaire de définir un thème concret autour duquel on pourrait « bâtir » les exercices. Pour le besoin du présent projet, nous avons choisi un thème qui se trouve à la frontière de la grammaire et du

lexique, à savoir la construction *avoir* N. Une esquisse théorique et la relation avec l'enseignement/apprentissage de la langue sont à trouver dans le chapitre 3.

Du point de vue méthodologique, nous avons collecté un « nouveau » corpus didactique qui se base sur les données du corpus parallèle InterCorp. La raison d'être du corpus didactique est le fait qu'InterCorp n'a pas d'ambitions didactiques et, par conséquent, les occurrences ont dû être triées. Le procédé entier est décrit dans le chapitre 4.

La partie la plus importante du travail est le chapitre 5, qui présente différents types d'exercices. En général, le but n'était pas de proposer un type général d'exercice qui est toujours employable ; au contraire, notre objectif était de présenter nos idées, nos propositions et d'inspirer et encourager les enseignants potentiellement intéressés. De plus, nous avons voulu montrer qu'il est parfaitement possible d'ajuster les activités aux buts désirés.

Ce qui s'est montré un autre avantage, c'est la possibilité d'employer le corpus pour n'importe quel type d'activité (grammaire, lexique, prononciation, révision...) et de recouvrir tout le processus d'éducation : présentation du thème (analyse et synthèse possibles), application, révision et test.

Pour la présentation du phénomène concret, on peut utiliser p. ex. un exercice d'introduction, appelé motivation dans notre texte, et ensuite l'analyse de mini-corpus dressé à partir du corpus didactique. L'application des connaissances acquises est possible p. ex. à l'aide des phrases à trous ou de la traduction. Le même pour la révision (ces deux étapes ont été confondues dans nos exercices) et pour le test final qui peut, lui aussi, être bâti à partir du corpus didactique et de la même méthodologie.

Comme nous l'avons mentionné dans la partie pratique, il est recommandé d'avoir toujours une « version prof » prête, c'est-à-dire un corrigé des activités préparées à partir du corpus. Bien qu'il s'agisse le plus souvent d'exercices avec plusieurs bonnes réponses, il est utile de posséder un guide qui peut, « au cas où », résoudre les hésitations potentielles.

Finalement, nous tenons à souligner ce qui a été déjà plusieurs fois dit : dans ce mémoire, nous n'avons pas ciblé les exercices les plus efficaces – l'objectif primaire était de proposer la typologie la plus vaste possible de diverses activités qui peuvent enrichir les cours du français langue étrangère. Ainsi, nous avons toujours visé, au détriment de l'efficacité, cet idéal « éventail de possibilités ».

7 Bibliographie

- Anderman, E. M., & Anderman, L. H. (2021). *Classroom motivation: linking research to teacher practice* (3^e éd.). Routledge, Taylor & Francis Group.
- André, V. (2020). Des corpus pour développer les compétences métalinguistiques des adultes francophones natifs en insécurité langagière. *Revue TDFLE*, (75).
- Aubin, S., & Bírová, J. (2019). De la valeur de la Philologie française en Europe aujourd’hui : Pensées philologique et didactique. *Synergies Europe*, 14(1), 7–20.
- Benko, V. (2014). Aranea: Yet Another Family of (Comparable) Web Corpora. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, Speech and Dialogue: 17th International Conference, TSD 2014* (pp. 257–264). Springer International Publishing Switzerland.
- Blanche-Benveniste, C. (1997). *Approches de la langue parlée en français*. Ophrys.
- Boulton, A., & Cobb, T. (2017). Corpus Use in Language Learning: A Meta-Analysis. *Language Learning*, 67(2), 348–393.
- Boulton, A. (2017). Corpora in Language Teaching and Learning. *Language Teaching*, 50(4), 483–506.
- Čermák, F. (2013). *Jazyk a jazykověda* (2^e éd.). Karolinum.
- Čermák, F. (2017a). *Korpus a korpusová lingvistika*. Karolinum.
- Čermák, F. (2017b). KORPUS. In: P. Karlík, M. Nekula & J. Pleskalová (Eds.), *CzechEncy – Nový encyklopedický slovník češtiny* [disponible en ligne].
- Čermák, F., & Rosen, A. (2012). The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics*, 17(3), 411–427.
- De Knop, S., & Gilquin, G. (Eds.). (2016). *Applied Construction Grammar*. De Gruyter.
- Detey, S., Durand, J., Laks, B., & Lyche, C. (2011). De l’enseignement des variétés avec corpus à la formation à la variation sur corpus : des outils pour le français parlé. In O. Bertrand & I. Schaffner (Eds.), *Variétés, variations et formes du français* (pp. 407–427). Editions de l’Ecole Polytechnique.
- Di Vito, S. (2013). L’utilisation des corpus dans l’analyse linguistique et dans l’apprentissage du FLE. *Linx*, (68–69), 159–176.
- Di Vito, S. (2019). La didattica della lingua francese, la linguistica dei corpora e l’inclusione. *Educazione Linguistica Language Education*, 8(2).
- Dohalská, M., & Schulzová, O. (2015). *Fonetika francouzštiny* (Čtvrté vydání). Univerzita Karlova v Praze, nakladatelství Karolinum.
- Dohalská, M., & Suková Vychopňová, K. (Eds.). (2015). *Didactique de la phonétique et phonétique en didactique du fle*. Université Charles de Prague, éditions Karolinum.
- Fenclová, M. (2010). *Stručná korektivní fonetika francouzštiny: výklady, přehledy a cvičení* (2^e éd.). Západočeská univerzita v Plzni.
- Flowerdew, J. (2009). Corpora in Language Teaching. In M. H. Long & C. J. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 327–350). Wiley-Blackwell.
- Flowerdew, J. (2009). Corpora in Language Teaching. In M. H. Long & C. J. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 327–350). Wiley-Blackwell.
- Forti, L. (2019). Evaluating Data-Driven Learning Effects in the Italian L2 Classroom. *Educazione Linguistica Language Education*, 8(2).
- Fried, M., & Östman, J.-O. (2004). Construction Grammar: A thumbnail sketch. In M. Fried & J.-O. Östman (Eds.), *Construction Grammar in a Cross-Language Perspective* (pp. 11–86). John Benjamins Publishing Company.

- Grevisse, M., & Goosse, A. (2016). *Le bon usage : Grevisse langue française* (7^e éd.). De Boeck Supérieur.
- Günday, R., Çakir, F., & Atmaca, H. (2017). Quelle démarche suivre en didactique de langue étrangère : la grammaire implicite ou la grammaire explicite ?. *HUMANITAS – Uluslararası Sosyal Bilimler Dergisi*, 247–262.
- Hendrich, J., Radina, O., & Tláškal, J. (2001). *Francouzská mluvnice* (3^e éd.). Fraus.
- Hoffmann, T., & Dancygier, B. (2017). From Constructions to Construction Grammars. In B. Dancygier (Ed.), *The Cambridge Handbook of Cognitive Linguistics* (pp. 284–309). Cambridge University Press.
- Chlumská, L. (2014). Není korpus jako korpus: Korpusy v kontrastivní lingvistice a translologii. *Časopis pro moderní filologii*, 96(2), 221–232.
- Choděra, R. (2013). *Didaktika cizích jazyků: úvod do vědního oboru*. Academia.
- Chun, S. (2016). High School Teachers' Need for DDL Materials: Appropriate Amount of Concordance Input. *Multimedia-Assisted Language Learning*, 19(1), 32–57.
- Kamber, A., & Dubois, M. (2016). Corpus, grammaire et français langue étrangère : une concordance nécessaire. *Linguistik Online*, 78(4).
- Le Cunff, C. (2005). De l'usage des corpus en didactique de l'oral : recherche et formation. *Rivages linguistiques*, 397–406.
- Le Foll, E. (Ed.). (2021). *Creating Corpus-Informed Materials for the English as a Foreign Language Classroom. A step-by-step guide for (trainee) teachers using online resources* (3^e éd.). Open Educational Resource.
- Leńko-Szymańska, A. (2017). Training teachers in data-driven learning: Tackling the challenge. *Language, Learning and Technology*, 21(3), 217–241.
- Mathieu, Y. Y. (2003). La Grammaire de Construction. *Linx*, (48), 43–56.
- McEnery, T., & Wilson, A. (2001). *Corpus linguistics: an introduction* (2^e éd.). Edinburgh University Press.
- Moon, S., & Oh, S. (2018). Unlearning Overgenerated "Be" through Data-Driven Learning in the Secondary EFL Classroom. *ReCALL*, 30(1), 48–67.
- Nádvorníková, O. (2017). Le corpus multilingue InterCorp : nouveaux paradigmes de recherche en linguistique contrastive et en traductologie. *Studii de Lingvistica*, 7, 67–88.
- Nádvorníková, O., Polická, A., Šotolová, J., & Vurm, P. (2010). Využití InterCorpu ve vysokoškolských kurzech francouzské filologie (motivace a cíle meziuniverzitní spolupráce romanistů při využití InterCorpu ve výuce). In *Mnohojazyčný korpus InterCorp: Možnosti studia* (pp. 232–240). NLN.
- Pořízka, P. (2014). *Tvorba korpusů a vytěžování jazykových dat: metody, modely, nástroje*. Vydavatelství Filozofické fakulty Univerzity Palackého.
- Radimský, J. (2010). *Verbo-nominální predikát s kategoriálním slovesem* [Habilitationní práce]. FF MU.
- Riegel, M., Pellat, J.-C., & Rioul, R. (2018). *Grammaire méthodique du français* (7^e éd.). Presses Universitaires de France.
- Saussure, F. de. (1996). *Kurs obecné lingvistiky* (2^e éd.). Academia.
- Skalková, J. (2007). *Obecná didaktika: vyučovací proces, učivo a jeho výběr, metody, organizační formy vyučování* (2^e éd.). Grada.
- Šabršula, J. (1986). *Vědecká mluvnice francouzštiny*. Academia.
- Šebesta, K., & Škodová, S. (Eds.). (2012). *Čeština – cílový jazyk a korpusy*. Technická univerzita v Liberci.

- Šebesta, K., & Šormová, K. (2019). *Korpusy v jazykovém vyučování II*. Filozofická fakulta Univerzity Karlovy.
- Vlasák, V. (2002). *Francouzsko-český, česko-francouzský slovník*. Leda.
- Vyatkina, N. (2016). Data-Driven Learning for Beginners: The Case of German Verb-Preposition Collocations. *ReCALL*, 28(2), 207–226.
- Владиславович, Д. А., Самуиловна, К. М., & Константиновна, В. Е. (2020). Теоретико-прикладное значение корпусов в компьютерной лингводидактике. *LITERA*, (1), 200–216.

Les corpus utilisés

- Benko, V. (2015). *Araneum Francogallicum Maius, verze 15.03*. Praha: Ústav Českého národního korpusu FF UK.
- Nádvorníková, O., & Vavřín, M. (2020). *Korpus InterCorp – francouzština, verze 13 z 1. 11. 2020*. Praha: Ústav Českého národního korpusu FF UK.
- Rosen, A., Vavřín, M., & Zasina, A. J. (2020). *Korpus InterCorp – čeština, verze 13 z 1. 11. 2020*. Praha: Ústav Českého národního korpusu FF UK.

Remarque : Tous les corpus cités sont disponible sur <http://www.korpus.cz>.

Les « programmes »

Dans le mémoire, nous avons cité deux documents ministériels portant sur l'enseignement – le programme « RVP pro gymnázia » (dans le texte comme « RVP ») et « RVP v oblasti informatiky a ICT ». Les deux documents sont disponible sur le site du Ministère de l'Éducation, de la Jeunesse et des Sports (« Ministerstvo školství, mládeže a tělovýchovy České republiky ») : <https://www.msmt.cz/vzdelavani/stredni-vzdelavani>.

8 Appendix

Remarque : Dans cette partie, nous présenterons les exercices décrits dans le texte précédent. Pour rendre nos activités encore plus claires, nous les avons pourvues des commentaires (en **police rouge**).

Tous les exercices sont censés être utiles dans l'enseignement secondaire (niveau lycée). Néanmoins, nous sommes sûr qu'il est possible de les employer p. ex. dans les cours pratiques pour les apprenants adultes. Dans ce qui suit, nous n'indiquerons donc que le niveau de langue.

Ex. 1

[Niveau A2–B1 ; à apprendre : la construction avoir N, conjugaison du verbe avoir]

Pour commencer tous les exercices présentés, on peut utiliser une activité de motivation (voir chap. 5.1.1). Ici, en considération du niveau, il est possible d'entamer le cours avec la conjugaison du verbe avoir.

Dans l'exercice suivant, il y a dix mini-corpus de quatre occurrences parallèles. Nous recommandons d'assigner une partie à un étudiant (ou un groupe d'étudiants) : chacun lit ses quatre phrases, les analyse et ensuite les présente au reste de la classe. Dans la première étape, chaque étudiant s'entraîne en analyse de la langue, dans la seconde, les étudiants apprennent les locutions sous forme de « vocabulaire ».

Remarque : lors de notre pilotage, les étudiants n'ont pas compris tout de suite ce qu'on veut qu'ils fassent. Ainsi, nous recommandons d'avoir un exemple prêt à l'aide duquel on peut expliquer le devoir – p. ex. avoir besoin de QC (forme), « potřebovat něco » (sens et la traduction).

1	<i>On a besoin de moi. Tu vas en avoir besoin. Je n'ai besoin de rien. J'ai besoin d'une cigarette.</i>	Potřebuji mě. Budeš to potřebovat. Nic mi nechybí. Potřebuju cigaretu.
2	<i>J'ai chaud. Tu dois avoir chaud. J'ai froid et j'ai chaud ! J'ai couru, j'ai chaud.</i>	Je mi horko. Musí ti být horko. Je mi horko i zima zároveň! Běžela jsem, je mi horko.
3	<i>Tu as envie de m'embrasser ? J'ai envie d'aller plus loin. J'ai envie de partir. J'ai envie de chocolat.</i>	Ted' mi chceš dát pusu, že jo? Chci pokračovat. Už chci jít. Dal bych si trochu čokolády.
4	<i>Je commence à avoir faim. Je suis épuisée et j'ai faim. Crois-tu qu'Alex a faim ? Bon, j'espère que vous avez faim.</i>	Začínám mít hlad. Jsem unavená a mám hlad. Myslíš, že má Alex hlad? Doufám, že máte hlad.
5	<i>J'ai froid. Tu as froid, non ? Tu sembles avoir froid.</i>	Je mi zima. Musí ti být zima, ne? Vypadáš, že je ti zima.

L'exercice B est un exercice « traditionnel » – il faut compléter les « composantes nominales » des locutions à apprendre. Au cas idéal, les étudiants doivent le faire sans regarder les conclusions tirées de la partie précédente. Sinon, il est possible de les consulter, si nécessaire. D'après notre expérience, la seconde variante est plus probable.

B) Complétez les phrases.

1. Tu dois avoir _____. (*Musíš mít hlad.*)
2. Je suis fatiguée, j'ai _____. (*Jsem unavená a je mi zima.*)
3. J'ai _____ de faire ça... (*Dělám chybu.*)
4. Je suppose que vous avez _____. (*Vsadím se, že si dáte něco k pití.*)
5. J'ai _____ de riz, de haricots, de lait en poudre, de bière et votre meilleur whisky. (*Pořebuju rýži, fazole, sušené mléko, pivo a tvoji nejlepší whisky.*)
6. Le spectacle d'ouverture a _____ ce soir. (*Zahajovací slavnost bude dnes večer.*)
7. Je commence à avoir _____. (*Začíná mi být zima.*)
8. Oui, j'imagine que vous avez _____. (*Jo, domnívám se že máte pravdu.*)
9. J'ai _____ partout. (*Všechno mě bolí.*)
10. Le procès aura _____ en Allemagne. (*Soud se bude konat v Německu.*)
11. Mère, père, vous aviez _____. (*Matko! Otče! Mýlili jste se!*)
12. Lucien avait _____ de pleurer. (*Lucien měl chuť dát se do pláče.*)
13. Je commence à avoir _____. (*Začínám cítit bolest.*)
14. Ils ont _____, ils ont _____. (*Je jim horko a mají hlad.*)

Jusqu'ici, on a entraîné la connaissance du système (et du lexique). Maintenant, nous voulons que les étudiants parlent et qu'ils emploient les locutions apprises de manière active. On les encourage de créer un dialogue en employant (pas forcément toutes) les phrases suivantes.

C) Créez un dialogue en utilisant les phrases suivantes :

- Tu as besoin de moi ?
- Moi aussi j'ai envie d'aller à la plage.
- J'ai mal à la lèvre !
- J'ai raison ?
- Bon Dieu, vous devez avoir soif.
- Tu as tort, tu sais.

L'activité suivante porte sur la traduction. Les traductions présentées dans le corrigé ne sont pas, bien évidemment, les seules possibles. Néanmoins, les étudiants doivent (en traduisant vers le français) toujours employer la construction avoir N.

D) Traduisez dans l'autre langue :

J'ai envie de me promener.

J'aime avoir chaud.

Je mange quand j'ai faim, je bois quand j'ai soif.
.....

J'ai besoin d'aide.

Moji holku bolí zuby.

Plete se.

Máš hlad?

Il sait que j'ai raison.

Ex. prof (corrigé de l'Ex. 1)

Ici, nous présentons la version « prof », un corrigé de l'Ex. 1.

1	<i>On a besoin de moi. Tu vas en avoir besoin. Je n'ai besoin de rien. J'ai besoin d'une cigarette.</i>	Potřebuji mě. Budeš to potřebovat. Nic mi nechybí. Potřebuju cigaretu.
2	<i>J'ai chaud. Tu dois avoir chaud. J'ai froid et j'ai chaud ! J'ai couru, j'ai chaud.</i>	Je mi horko. Musí ti být horko. Je mi horko i zima zároveň! Běžela jsem, je mi horko.
3	<i>Tu as envie de m'embrasser ? J'ai envie d'aller plus loin. J'ai envie de partir. J'ai envie de chocolat.</i>	Ted' mi chceš dát pusu, že jo? Chci pokračovat. Už chci jít. Dal bych si trochu čokolády.
4	<i>Je commence à avoir faim. Je suis épuisée et j'ai faim. Crois-tu qu'Alex a faim ? Bon, j'espère que vous avez faim.</i>	Začínám mít hlad. Jsem unavená a mám hlad. Myslíš, že má Alex hlad? Doufám, že máte hlad.
5	<i>J'ai froid. Tu as froid, non ? Tu sembles avoir froid. J'ai mal à la jambe, j'ai froid et j'ai faim.</i>	Je mi zima. Musí ti být zima, ne? Vypadáš, že je ti zima. Mám zraněnou nohu, je mi zima a mám hlad.
6	<i>La rencontre aura lieu ce vendredi 13. Quand a lieu cette course ? Le mariage a lieu samedi, nous sommes jeudi. Le crime a eu lieu ici.</i>	Schůzka bude třináctého v pátek tento měsíc. Kdy je ten závod? Svatba je v sobotu, dnes je čtvrtek. Zločin se stal tady.
7	<i>J'ai mal à la gorge. J'ai mal à la tête. Tu as mal ? J'ai mal au ventre.</i>	Bolí mě krk. Mám bolení hlavy. Máš bolesti? Je mi špatně od žaludku.
8	<i>Tu as raison, mais... Bien sûr que j'ai raison. Ma mère avait raison. Ouais, tu as raison. Excuse-moi.</i>	To je pravda, ale... Jasně že mám pravdu. Máma měla pravdu. Jo, máte pravdu. Máte pravdu. Promiňte.
9	<i>Tu as soif ? Institutrice, j'ai soif. Tout le monde a soif ? J'ai chaud, j'ai soif !</i>	Máš žízeň? Paní učitelko, mám žízeň. Máte všichni žízeň? Je mi horko, mám žízeň.
10	<i>On avait tort.</i>	Zmýlili jsme se.

J'avais tort.

Je sais que j'avais tort et je suis désolé.

T'as tort.

Mýlil jsem se.

Vím, že jsem udělal špatnou věc a omlouvám se.

Jdeš na to špatně, chlape.

Devoirs :

1. Lisez les phrases.
2. Décrivez la construction (sens, complémentation).
3. Présentez-la aux autres !

A) Classez les constructions d'après leurs compléments (prépositions). Indiquez la traduction en tchèque.

Il est possible d'écrire cet exercice au tableau.

À	De	Temps	Aucun complément
MAL (MÍT BOLESTI)	BESOIN (POTREBOVAT) ENVIE (CHTÍT, MÍT CHUŤ)	LIEU (KONAT SE)	CHAUD (BÝT HORKO) FAIM (MÍT HLAD) FROID (BÝT ZIMA) MAL (BOLET) RAISON (MÍT PRAVDU) SOIF (MÍT ŽÍZEŇ) TORT (MÝLIT SE)

B) Complétez les phrases.

1. Tu dois avoir FAIM. (*Musíš mít hlad.*)
2. Je suis fatiguée, j'ai FROID. (*Jsem unavená a je mi zima.*)
3. J'ai TORT de faire ça... (*Dělám chybu.*)
4. Je suppose que vous avez SOIF. (*Vsadím se, že si dáte něco k pití.*)
5. J'ai BESOIN de riz, de haricots, de lait en poudre, de bière et votre meilleur whisky. (*Pořebuju rýži, fazole, sušené mléko, pivo a tvoji nejlepší whisky.*)
6. Le spectacle d'ouverture a LIEU ce soir. (*Zahajovací slavnost bude dnes večer.*)
7. Je commence à avoir FROID. (*Začíná mi být zima.*)
8. Oui, j'imagine que vous avez RAISON. (*Jo, domnívám se že máte pravdu.*)
9. J'ai MAL partout. (*Všechno mě bolí.*)
10. Le procès aura LIEU en Allemagne. (*Soud se bude konat v Německu.*)
11. Mère, père, vous aviez TORT. (*Matko! Otče! Mýlili jste se!*)
12. Lucien avait ENVIE de pleurer. (*Lucien měl chuť dát se do pláče.*)
13. Je commence à avoir MAL. (*Začínám cítit bolest.*)
14. Ils ont FAIM, ils ont CHAUD. (*Je jim horko a mají hlad.*)

C) Créez un dialogue en utilisant les phrases suivantes :

TRADUCTION RECOMMANDÉE AVANT LA CRÉATION DES DIALOGUES.

- Tu as besoin de moi ? – NEPOTŘEBOVAL BYS NĚCO?
- Moi aussi j'ai envie d'aller à la plage. – TAKY BYCH JEL K MOŘI.
- J'ai mal à la lèvre ! – BOLÍ MĚ RET.
- J'ai raison ? – NEMÁM PRAVDU?
- Bon Dieu, vous devez avoir soif. – PROBOHA, TO MUSÍTE MÍT ŽÍZEŇ.
- Tu as tort, tu sais. – MÝLÍŠ SE, VÍŠ?

D) Traduisez dans l'autre langue :

J'ai envie de me promener. – ASI SE PŮJDU PROJÍT.

J'aime avoir chaud. – MÁM RÁD TEPLŮ.
 Je mange quand j'ai faim, je bois quand j'ai soif. – VÍM, KDY MÁM ŽÍZENĚ, JÍM, KDYŽ MÁM HLAD.
 J'ai besoin d'aide. – MUSÍŠ MI POMOCT.
 MON AMIE A MAL AUX DENTS. – Moji holku bolí zuby.
 IL A TORT. – Plete se.
 T'AS FAIM? – Máš hlad?
 Il sait que j'ai raison. – VÍ, ŽE MÁM RECHT.

Ex. 2

[Niveau B1 ; à apprendre : la construction avoir N, conjugaison du verbe avoir]

Dans ce qui suit, chaque étudiant essaie de traduire les phrases vers le tchèque. Puis, il est possible de comparer les traductions en classe, et ensuite consulter les traductions proposées par le corpus. Les étudiants sont invités de relier les phrases avec leurs équivalents tchèques ce qui est une étape facile, étant donné les traductions déjà faites. Pour que cette étape ne soit pas si facile, nous y avons mis un « intrus », une phrase qui n'a pas d'équivalent.

Essayez de traduire ces phrases :

1. *Vous avez droit à un avocat.*
2. *Tu as rendez-vous à 5h pour tes cheveux.*
3. *Pas de quoi avoir honte.*
4. *On a affaire à la mafia chinoise.*
5. *J'ai honte de ce qu'on a fait.*
6. *Il a accès à tes armes ?*
7. *J'ai rendez-vous avec Mme Devereaux.*
8. *Tu as peur du gouvernement ?*
9. *J'ai droit à une explication.*
10. *Outre Pierre, qui a accès à ces documents ?*
11. *Maintenant nous savons à quoi nous avons affaire.*
12. *Docteur, j'ai peur.*

Reliez les phrases aux traductions :

- A. *Doktore, já mám strach.*
- B. *Dlužíte mi vysvětlení.*
- C. *Kdo jiný, kromě Petra, má k těm dokumentům přístup?*
- D. *Bojíš se vlády?*
- E. *Máte právo na advokáta.*
- F. *Dobrá, tak teď víme, proti čemu bojujeme.*
- G. *Máme tu čest s čínskou mafii.*
- H. *Má přístup k vašim pistolím?*
- I. *Moc se stydím za to, co jsme provedli.*
- J. *O páté máš schůzku s kadeřnicí.*
- K. *Paní Devereauxová mě čeká.*
- L. *Není se za co stydět.*

Correspondances : 1 ..., 2 ..., 3 ..., 4 ..., 5 ..., 6 ..., 7 ..., 8 ..., 9 ..., 10 ..., 11 ..., 12 ...

A) Traduisez les constructions à avoir en tchèque et indiquez la complémentation.

(p. ex. avoir besoin de = *potřebovat něco*)

avoir d _____ = _____
avoir r _____ = _____
avoir h _____ = _____
avoir af _____ = _____
avoir ac _____ = _____
avoir p _____ = _____

B) Cherchez le/s mot/s omi/s.

1. On a _____ à un monstre. (*Věřte mi, tenhle chlap je netvor.*)
2. _____ avait _____. (*Měli jsme schůzku.*)
3. J'avais _____. (*Bál jsem se.*)
4. Donc, _____ avez accès _____ décisions, militaires ou politiques. (*Takže máte přístup ke všem přijímaným rozhodnutím, vojenským i politickým?*)
5. Tu devrais _____ honte ! (*Styd' se!*)
6. Vous _____ silence. (*Máte právo nevypovídat.*)

La conception de l'activité suivante est libre. Néanmoins, nous recommandons la forme de discussion dirigée où les étudiants font deux groupes (« pour » et « contre »).

C) Est-ce que vous y avez droit ? Discutez-en !

Tout le monde a droit à son avis.

La conception de l'activité suivante est libre. L'enseignant peut définir la forme de texte ou non. Nous ne l'avons pas définie, car nous croyons que la phrase d'introduction, qui est obligatoire, mènera les étudiants vers une réflexion.

D) Écrivez un petit texte. Employez la phrase d'introduction et les constructions suivantes :

avoir droit, avoir peur, avoir honte, avoir accès, avoir rendez-vous

Je crois que tout le monde a droit...

Cet exercice « structural » porte sur un phénomène grammatical concret.

E) Complétez les formes du COD et COI.

1. Il ____ a peur évidemment. (*Je z toho očividně vystrašený.*)
2. Qui ____ avait accès ? (*Kdo k tomu měl přístup?*)
3. J'ai honte de ce qu'on a fait. (*Moc se stydím za to, co jsme provedli.*) = J' ____ ai honte.
4. Ils ____ ont droit. (*Zaslouží si to.*)
5. Il était l'un des 3 hommes à avoir accès à l'ascenseur. (*Byl jeden ze tří lidí, co měli přístup do toho výtahu.*) = Il était l'un des 3 hommes à ____ avoir accès.

De nouveau, nous ajoutons une activité portant sur l'oral (les consignes voir *supra*). Pour ce qui concerne les traductions (phrases tchèques), c'est le choix de l'enseignant s'il veut les ajouter ou pas. Nous les avons ajoutées pour indiquer cette possibilité, mais nous croyons qu'elles sont (parfois) superflues. Il est raisonnable de les ajouter au cas de phrases « compliquées », difficiles.

F) Créez les dialogues en employant les phrases suivantes :

- On a accès à la boîte mail de Ross. (Nabořili jsme se Rossovi do mailu.)
- Quand j'ai vu le sang... j'ai eu peur. (Když jsem viděla krev, byla jsem vyděšená.)
- Comme j'ai honte ! (Já se tak stydím.)
- Tu as peur de quoi ? (Čeho se bojíš?)
- J'ai rendez-vous demain, pour un boulot. (Zítra jdu na pohovor. Kvůli práci.)

G) Traduisez dans l'autre langue.

1. Ano, mám schůzku.
2. Parce qu'ils ont peur de tomber et de mourir.
3. Ici, tout le monde a droit à son pourboire, c'est la loi !
4. Ty se stydíš?
5. Tu as honte de travailler avec nous ?
6. Vous ignorez à qui vous avez affaire.
7. Je očividně vystrašený.

Ex. 3

[Niveau B1 ; à apprendre : la construction *avoir* N, conjugaison du verbe *avoir*, traits principaux du français parlé]

Dans cette partie, nous présenterons les diapositives PowerPoint. Nous n'ajoutons que quelques diapositives qui, dans la présentation entière, se répètent (p. ex. la diapositive avec le corrigé suit la diapositive sans corrigé...).

La diapositive suivante contient plusieurs activités : d'abord, les étudiants sont invités à traduire les phrases françaises, puis nous leur avons montré les équivalents du corpus (dans l'ordre aléatoire) qui sont à relier avec les phrases françaises. Ensuite vient le corrigé, et finalement, les étudiants ont été encouragés à trouver (d'après les deux versions – française et tchèque) les traits de la langue parlée qui, pour contrôler, ont été marqué avec une couleur différente.

Cherchez les équivalents

- | | |
|--------------------------------------|---------------------------------|
| 1. J'ai chaud. | A. Mám rád teplo. |
| 2. J'aime avoir chaud. | B. Páni, to bylo o chlup, kámo. |
| 3. T'as eu chaud. | C. Je mi horko. |
| 4. On a eu chaud, mon vieux . | D. To bylo o fous. |

Corrigé : 1C, 2A, 3D, 4B

Avoir chaud = být (někomu) horko

TRADUISEZ :

- Musí ti být horko.
- Je ti horko?
- Je mi horko i zima zároveň!
- J'ai couru, j'ai chaud.
- On a eu chaud !

Voici le corrigé de la diapositive précédente. Ici aussi, les étudiants devaient chercher les traits de la langue parlée/familiale (rouges).

Avoir chaud = být (někomu) horko

TRADUISEZ :

- Musí ti být horko. = *Tu dois avoir chaud.*
- Je ti horko? = *Tu as chaud, **toi** ?*
- Je mi horko i zima zároveň! = *J'ai froid et j'ai chaud !*
- J'ai couru, j'ai chaud. = *Běžela jsem, je mi horko.*
- **On a eu chaud !** = *J'en ai eu assez!*

Ex. 4 (test pour l'Ex. 1)

Ce test, créé à l'instar des activités présentées ci-dessus, contrôle la maîtrise du phénomène ciblé – la construction avoir N.

A) Complétez les phrases : [5 p.]

Tu dois avoir _____. (*Musíš mít hlad.*)

J'ai _____, j'ai _____ ! (*Je mi horko, mám žízeň.*)

J'ai _____ de partir. (*Už chci jít.*)

Oh, ce que j'ai _____ ... (*Je mi strašný horko.*)

B) Traduisez : [10 p.]

Potřebuju cigaretu.

Mýlil jsem se.

J'ai envie de chocolat.

Mám bolení hlavy.

Ma mère avait raison.

C) Liez les phrases : [5 p.]

- | | |
|-----------------------------------|---------------------------------------|
| 1. J'ai mal partout. | A. Potřebují mě. |
| 2. Je suppose que vous avez soif. | B. Začínám mít hlad. |
| 3. Il a raison, tu as tort. | C. Všechno mě bolí. |
| 4. On a besoin de moi. | D. Vsadím se, že si dáte něco k pití. |
| 5. Je commence à avoir faim. | E. Zločin se stal tady. |

1 __, 2 __, 3 __, 4 __, 5 __

__ p. en somme

Le test porte uniquement sur la lecture et/ou l'écrit. Si nécessaire, il serait possible de tester également l'oral – voir chap. 5.1.10.