

## L'UTILISATION DES CONJONCTIONS COMME OUTIL DE COHÉSION TEXTUELLE DANS LE TCHÈQUE DE LOCUTEURS NON NATIFS

Kateřina RYSOVÁ, Magdaléna RYSOVÁ, Eva HAJIČOVÁ  
Université Charles, Prague (MFF UK)

**Abstract (En):** In the paper, we explore cohesive devices in Czech texts written by students – non-native speakers. Specifically, we focus on conjunctions (significantly contributing to text coherence) from the point of view of text genres; we analyse four basic genres, namely narration, information, argumentation and description, and we follow the relative occurrence of text conjunctions across them. Methodologically, we use data obtained from the CzeSL-SGT corpus (AKCES 5; Šebesta et al., 2014) containing texts written by non-native speakers of Czech. The texts are annotated with text genres, students' age, text evaluation (A–C levels according to CEFR, the Common European Framework of Reference for Languages) and other additional information. For our analysis, we selected texts of narration, information, argumentation and description written by students at the age of 16+. We examined the occurrence of conjunctions at the individual CEFR levels (A, B and C). The absolute numbers of occurrences were converted to relative frequencies (the i.p.m. value – Instances per Million Positions). The aim of the analysis is to find out whether the particular genres differ in frequency and diversity of text conjunctions and whether some difference also occurs across the CEFR levels – A: basic user, B: independent user, C: proficient user. Based on the corpus data, we conclude that conjunction devices contributing to text coherence occur most frequently in the genre of argumentation, followed by narration, less in description and information. The individual genres are explored in detail also with respect to the CEFR levels of A–C. A and B levels exhibit the same tendencies, i.e. conjunction devices occurred most frequently in genres with the following order: argumentation (A: 90,680.64 i.p.m., B: 90,088.53 i.p.m.), narration (A: 72,287.64 i.p.m., B: 77,511.01 i.p.m.), description (A: 66,940.08 i.p.m., B: 66,856.51 i.p.m.), and information (A: 51,424.80, B: 61,542.47). In the C category, the relative occurrence of conjunctions in narration (83,751.65 i.p.m.) and argumentation (83,376.42 i.p.m.) appeared to be similar, whereas they were less frequent in information (66,648.96 i.p.m.) and description (57,153.67 i.p.m.). The results, however, may be influenced by the lower number of data in the C category. The results demonstrated that the occurrence of conjunction devices is related to the nature of the text genres: argumentation and narration appear to be more higher-structured genres in terms of discourse relations. They contain more complex coherence relations than description or information, and therefore, they also contain more conjunction devices that make discourse relations better comprehensible and prevent misunderstanding of the text as a whole.

**Keywords (En):** discourse connectives; text genres; foreign language acquisition; non-native speakers; Czech

**Mots-clés (Fr):** connecteurs du discours; typologie textuelle; acquisition d'une langue étrangère; locuteurs non natifs; Tchèque

**DOI :** 10.32725/eer.2022.005

## **Introduction**

Le texte, en tant qu'enregistrement d'une activité langagière, constitue une des formes élémentaires de la communication linguistique.<sup>1</sup> Par son biais, un auteur partage un contenu à un destinataire. Un texte, qui est fonctionnel du point de vue communicatif, est une structure complexe – plusieurs niveaux linguistiques s'y retrouvent. En même temps, il est divisible en unités plus petites, comme des paragraphes, des phrases complexes, des propositions ou encore des locutions. Ces unités sont liées entre elles par des liens fonctionnels.

DANEŠ (1989) détermine dans le texte cinq niveaux de structuration : (1) le niveau des relations sémantiques interpropositionnelles reflétant les relations entre les objets et les situations du monde réel, (2) celui des relations isotopiques concernant l'identité des objets de la parole, (3) celui des relations thématiques – structure informationnelle de la phrase et progressions thématiques (ces relations sont partiellement liées au réseau des relations isotopiques), (4) celui des relations compositionnelles déterminant les liens entre les unités textuelles (soit des fonctions structurant dans un plan d'ensemble les parties informatives individuelles du texte) et (5) celui des structures de l'interaction performative se rapportant à la fonction énonciative et à l'intention. Daneš considère le niveau des relations sémantiques propositionnelles (1) comme extra-textuel, les niveaux des relations isotopiques (2), des relations thématiques (3) et des relations compositionnelles (4) comme purement intratextuels et la structure de l'interaction performative (5) comme un niveau visant à agir sur l'état de choses, à l'extérieur du texte.

Dans cet article, nous nous concentrons sur les relations internes du texte, en particulier sur les relations entre les unités textuelles que Daneš caractérise comme relations compositionnelles. L'objectif de cette recherche est d'étudier comment ces relations se manifestent dans un texte et quelles sont leurs fréquences dans différents types de textes.

Après un passage en revue des concepts élémentaires de la linguistique textuelle (partie 1), nous présenterons le corpus linguistique que nous avons exploité (partie 2). Nous aborderons ensuite la méthodologie (partie 3) et enfin l'analyse des conjonctions se manifestant dans les types de textes sélectionnés avec les résultats de nos recherches (partie 4).

### **1. Cohésion et cohérence textuelles**

Dans les paragraphes suivants, nous nous focaliserons sur plusieurs aspects de la linguistique textuelle. Nous présenterons d'abord les concepts clés, c'est-à-dire les approches propositionnelles et communicationnelles (1.1). Nous nous concentrerons en particulier sur les procédés de la cohésion textuelle et sur la définition des connecteurs comme balises élémentaires des relations sémantiques intratextuelles (1.2). Dans le cadre de l'analyse des connecteurs, nous nous pencherons sur la description de la partie du discours qui s'y trouve la mieux

---

<sup>1</sup> Ce travail est soutenu par le Ministère tchèque de l'éducation, de la jeunesse et de l'éducation physique de République tchèque, projet n° LM2018101 LINDAT/CLARIAH-CZ, et par la Fondation pour la science de la République tchèque (GA ČR) dans le cadre du programme EXPRO, projet n° GX20-16819X LUSyD.

représentée, à savoir les conjonctions (1.3). Nous analyserons ensuite les genres communicationnels (1.4) et les caractéristiques des procédés compositionnels, concrètement les procédés informatif, narratif, descriptif et explicatif (1.5).

### **1.1. Approches propositionnelle et communicative du texte**

La linguistique propose deux grandes approches pour l'analyse de texte (VIEHWEGER 1977 ; SGALL 1980 et autres). La première, dite communicationnelle, se concentre sur l'intentionnalité du texte. Celle-ci se base sur l'idée que le texte est construit « du haut vers le bas » et que son interprétation doit procéder dans le même sens. Cette approche envisage le texte comme un outil permettant aux êtres humains d'agir. Son analyse commence dès lors au niveau de l'intention de l'auteur et se poursuit vers les énoncés et les relations qu'ils entretiennent entre eux.

La deuxième approche, dite propositionnelle, se focalise sur le système linguistique. Elle part de l'hypothèse selon laquelle le texte est construit « du bas vers le haut » et son interprétation suit ce sens. L'analyse de texte démarre dès lors des unités linguistiques élémentaires, qui peuvent se combiner vers des ensembles plus grands que sont les paragraphes ou les chapitres. Il en résulte que cette analyse se concentre sur les relations à l'intérieur du texte et sur son organisation, c'est-à-dire sur sa composition, sa cohésion et sa cohérence occupant le centre de l'attention. Le présent article étudie les relations sémantiques interpropositionnelles et les moyens qui lient les unités linguistiques dans un texte cohésif et cohérent.

### **1.2. Les outils de la cohésion textuelle – les connecteurs du discours**

Du point de vue de l'approche propositionnelle, la cohésion et la cohérence représentent les propriétés fondamentales du texte. La cohérence concerne les relations de sens (sémantique et pragmatique), et la cohésion les relations formelles. En d'autres termes, la cohésion résulte des relations entre les moyens linguistiques dans la structure superficielle du texte, tandis que la cohérence relève des relations textuelles reposant sur la cohésion (RICKHEIT & SCHADE 2000). Lors de la description d'un texte cohésif et cohérent, il est dès lors important d'examiner les connexions des structures phrastiques avec les structures supra-phraseologiques et textuelles (HAJIČOVÁ 1993, 2012).

D'après ŠTÍCHA (2013), les relations textuelles élémentaires proviennent (1) des références à la situation de communication et aux coréférences, (2) des progressions thématiques, (3) des relations grammaticales entre les énoncés et (4) des connecteurs.

C'est cette dernière catégorie des outils de cohésion, les connecteurs du discours, qui nous intéresse pour le présent travail. Il s'agit d'outils linguistiques exprimant des relations entre des unités textuelles. Les connecteurs peuvent ainsi mettre en relation des propositions pour produire des phrases complexes (voir exemple 1). Ils peuvent également exprimer des relations inter-phraseologiques, et ainsi mettre en relation des phrases à l'intérieur d'un texte (voir exemple 2 – les exemples proviennent du corpus *Pražského závislostního korpusu*, HAJIČ *et al.* 2020).

(1) *Skibby ujel všem a vyhrál.*<sup>2</sup>

(2) *Prosperuje. A je to příklad pro všechny.*<sup>3</sup>

Différentes démarches pour délimiter les connecteurs du discours se distinguent les unes des autres. Dans certaines approches, sont inclus parmi les connecteurs des conjonctions, des syntagmes prépositionnels discursifs et des adverbes (PRASAD *et al.* 2008, 2010 et autres). SCHIFFRIN (1987) ou AIJMER (2002) quant à eux entendent la classe des connecteurs de manière plus large et y incluent des particules discursives et des groupes nominaux. FRASER (1999) considère comme connecteurs certaines locutions, comme *all things considered*. RYSOVÁ & RYSOVÁ (2014, 2018) partagent les connecteurs entre types primaires (comme *a, ale, proto*<sup>4</sup>) et secondaires (comme *z tohoto důvodu, za těchto podmínek*<sup>5</sup>) en fonction de leur degré de grammaticalisation.

La plupart des auteurs s'accordent sur le fait que les connecteurs du discours constituent des unités linguistiques capables d'exprimer des relations sémantiques entre deux portions de texte. Du point de vue des parties du discours, les conjonctions représentent des connecteurs typiques. C'est précisément sur les conjonctions que cette étude se focalise puisqu'elles sont les représentantes élémentaires et les plus nombreuses de la classe des connecteurs du discours.

### 1.3. Les conjonctions

Du point de vue des parties du discours, l'ensemble des conjonctions représente le sous-groupe le plus important de connecteurs du discours. Il s'agit de mots synsémantiques non flexionnels que la grammaire traditionnelle subdivise selon leur fonction syntaxique en conjonctions de coordination et de subordination. D'après KARLÍK (2016), les conjonctions de coordination lient des unités de différents types, comme des phrases, des syntagmes ou des mots. Les conjonctions de subordination ne le permettent pas. Exemples : *Petr a Pavel* vs. *\*Petr že Pavel*<sup>6</sup> (une conjonction de subordination entre des portions de texte autres que des phrases relève, selon P. Karlík, de l'ellipse : *Říkal, že já.*<sup>7</sup>). Les conjonctions de subordination sont en outre considérées comme des éléments à part entière des phrases où ils se trouvent, typiquement au début de celles-ci, tandis que les conjonctions de coordination n'appartiennent à aucune des unités textuelles qu'elles lient : *Spí, nebo bdí – \*Nebo bdí, spí* vs. *Tuší, že omdlí – Že omdlí, tuší.*<sup>8</sup>

Bien que les conjonctions constituent l'essentiel des connecteurs du discours, toutes n'en sont pas. Par exemple, les conjonctions introduisant une proposition complétive, comme *že*, ne sont pas considérées comme des connecteurs du discours car elles ne spécifient pas la nature du lien logico-sémantique (comme la cause, le but, la conséquence, etc.) qu'elles établissent entre les énoncés.

---

<sup>2</sup> *Skibby a dépassé tout le monde et il a gagné.* (N. d. T.)

<sup>3</sup> *Il prospère. Et c'est un exemple pour tous.* (N. d. T.)

<sup>4</sup> *et, mais, c'est pourquoi.* (N. d. T.)

<sup>5</sup> *raison pour laquelle, à ces conditions.* (N. d. T.)

<sup>6</sup> *Petr et Pavel vs \*Petr que Pavel* (N. d. T.)

<sup>7</sup> *Il a dit que moi.* (N. d. T.)

<sup>8</sup> *Il dort ou il veille – \*Ou il veille, il dort vs Il sent qu'il s'évanouit. – Qu'il s'évanouit, il le sent.* (N. d. T.)

Dans cet article, nous analyserons les conjonctions en tant qu'outil fondamental pour la production d'un texte cohésif et cohérent. L'étude s'effectuera sur différents types de textes relevant de différents genres compositionnels que nous présenterons plus en détail dans les parties suivantes.

#### **1.4. Les genres communicationnels**

LUCKMANN (1986) posa dans les années 80 du siècle dernier les fondements de la théorie des genres communicationnels que linguistes et sociologues approfondirent par la suite. Les genres communicationnels sont définis historiquement et culturellement, ils sont partiellement préconstruits et peuvent être réemployés dans des situations de communication récurrentes (GÜNTNER et KNOBLAUCH 1995). Ils fonctionnent ainsi comme un cadre pour la production et la réception d'un texte. En effet, les auteurs et les destinataires partagent un fond commun social et culturel, et, partant, des attentes spécifiques quant à la forme que doit revêtir un texte dans une situation de communication donnée. Les genres communicationnels peuvent donc être compris comme des modèles communicationnels.

Les textes d'un même genre communicationnel peuvent s'apparenter par des éléments de leur contenu (thèmes, motifs), par leur construction, par la variété de langue, le champ lexical, les outils stylistiques ou grammaticaux. Le propos de cet article est d'identifier ce qui distingue différents genres communicationnels du point de vue de leur construction, en particulier dans l'étendue et le mode d'utilisation des conjonctions.

#### **1.5. Les procédés compositionnels**

Les procédés compositionnels déterminent fondamentalement la construction d'un texte. Lors de la production d'un texte, l'auteur choisit ceux qu'il va utiliser pour exprimer le contenu visé selon la situation de communication, le contenu et le support formel. Plusieurs modèles (types) compositionnels spécifiques pour la production textuelle peuvent être identifiés en regard de ces choix.

KRČMOVÁ (2016) identifie les principaux procédés compositionnels suivants : (1) le procédé informatif, (2) le procédé narratif, (3) le procédé descriptif, (4) le procédé argumentatif, (5) le procédé explicatif. Selon J. V. BEČKA (1992 : 36), chacun de ces procédés « se distingue essentiellement par l'orientation des choix des outils lexicaux et de leurs fréquences (substantifs, verbes, adjectifs, concrets et abstraits, termes appréciatifs) ainsi que le choix des types de phrases et la manière dont elles s'agencent au niveau supraphrastique ».

MISTRÍK (1997) caractérise chacun de ces procédés sur la base de cinq couples de traits distinctifs : (1) la cohésion ou la non-cohésion, (2) le caractère explicatif ou énumératif, (3) la commutabilité ou successivité, (4) l'actualité ou la généralité, (5) la subjectivité ou l'objectivité. Le tableau 1 synthétise l'interprétation des différents types d'après Mistrík. À la différence de Krčmová, Mistrík distingue en outre le procédé dialogique. En revanche, il ne différencie pas le procédé argumentatif de l'explicatif et ne conserve que ce dernier terme. La présente étude s'intéresse en particulier au trait de cohésion / non-cohésion.

<b>Procédé compositionnel</b>		<b>Informatif</b>	<b>Narratif</b>	<b>Descriptif</b>	<b>Explicatif</b>	<b>Dialogique</b>
Cohésion	du texte	-	+	-	+	-
Non-cohésion		+	-	+	-	+
Caractère explicatif	dans le texte	-	+	-	+	+
Caractère énumératif		+	-	+	-	-
Commutabilité	dans le texte	+	-	+	-	-
Successivité		-	+	-	+	+
Actualité	du texte	+	+	-	-	+
Généralité		-	-	+	+	-
Subjectivité	du texte	-	+	+	-	+
Objectivité		+	-	-	+	-

*Tableau 1. Caractéristiques des types compositionnels d'après MISTRÍK (1997)*

Nous avons sélectionné pour cette recherche des textes rédigés en tchèque par des locuteurs non natifs dans des situations de communication équivalentes (textes monologiques écrits, compositions scolaires). Ils se distinguent néanmoins de par leur construction, c'est-à-dire par la conception de leur composition. Nous avons travaillé avec quatre ensembles de textes relevant des procédés informatifs, narratifs, descriptifs et argumentatifs. Nous analyserons ces procédés compositionnels au niveau de leurs caractéristiques de cohésion / non-cohésion.

MISTRÍK (1997) considère le procédé informatif comme non cohésif, et le plus simple du point de vue compositionnel. Il permet de présenter les faits, sans insister sur les relations qu'ils entretiennent entre eux, raison pour laquelle, dans les textes relevant de ce procédé compositionnel, les expressions assurant la cohésion du texte manquent généralement. « Dans le texte informatif, ni la cohésion, ni la fluidité, ni la continuité des faits ne sont pertinentes puisque chacun d'entre eux se suffit à lui-même sans égard à la construction du texte dans lequel il apparaît » (MISTRÍK 1997 : 343).

Le procédé narratif, à la différence du précédent, est considéré comme cohésif (MISTRÍK 1997). Ce sont les connecteurs qui assurent cette cohésion. Les textes de ce type contiennent par ailleurs de nombreux verbes. Le récit possède une temporalité précise et est généralement construit avec des relations de cause à effet. Nous pouvons donc nous attendre à y rencontrer, entre autres, des connecteurs temporels et de causalité.

MISTRÍK (1997) considère le procédé descriptif comme non cohésif. En cela il se rapproche du procédé informatif, « cependant, la non-cohésion des textes informatifs se situe dans la construction interne des phrases, tandis que la non-cohésion des textes descriptifs concerne les composants thématiques du texte. » (MISTRÍK 1997 : 387) MISTRÍK (1997 : 352) précise en outre qu'« il s'agit de non-cohésion contextuelle, pas au niveau des phrases complexes ».

Enfin, MISTRÍK (1997) considère le procédé explicatif (dans lequel il inclut le procédé argumentatif) comme cohésif. Pour la rédaction de ce type de texte, les auteurs expriment souvent, grâce à des connecteurs du discours, les relations entre les contenus présentés. Ce procédé est qualifié de plus cohésif parmi les différents types abordés.

## **2. Données linguistiques : le corpus AKCES 5 / CzeSL-SGT**

Pour la partie analytique de cette étude, nous avons exploité les données du corpus Czech as a Second Language with Spelling Grammar and Tags (AKCES 5 / CzeSL-SGT, ŠEBESTA et al. 2014). Il s'agit d'un corpus d'étudiants compilant les transcriptions de productions écrites en tchèque de non-natifs. Au total, 8 617 textes sont disponibles (111 000 phrases, 1 million de mots, 1,1 million de tokens) rédigés par 1 965 auteurs différents de 54 langues maternelles différentes. Les phénomènes linguistiques ont été annotés automatiquement. L'annotation recouvre le marquage des parties du discours, des catégories morphologiques et des lemmes. Des métadonnées accompagnent les textes et les pourvoient de données sur les auteurs et le type de texte, avec notamment les procédés compositionnels dominants.

## **3. Méthodologie**

Notre étude vise à déterminer si les différents types de textes se distinguent par le nombre ou par le type de conjonctions utilisées (les connecteurs de discours n'étant pas annotés dans le corpus, nous nous sommes concentrés sur les conjonctions qui, en tant que partie du discours, représentent la plupart des connecteurs du discours).

L'exploitation du corpus CzeSL-SGT a consisté à relever le nombre de conjonctions dans les textes de types argumentatifs, narratifs, descriptifs et informatifs. Nous avons également trié les textes selon les niveaux du Cadre européen commun de référence pour les langues (CEFR, 2011)<sup>9</sup>, à savoir : utilisateur débutant (A), utilisateur indépendant (B) et utilisateur avancé (C).

Le nombre de conjonctions identifiées est exprimé sous deux formes : le nombre absolu d'occurrences et la fréquence relative en i.p.m. (indice par million). La valeur de l'i.p.m. exprime le nombre moyen d'occurrences qui apparaîtraient dans un corpus hypothétique équivalent d'une longueur d'un million de mots.

Pour chaque type de texte, nous avons ensuite recherché les 10 conjonctions les plus fréquentes et comparé les tableaux ainsi obtenus.

---

<sup>9</sup> CEFR (l'acronyme en français) dans la suite du texte (N. d. T.).

#### 4. Résultats

L'objectif de cette étude était de comparer les nombres de conjonctions utilisées dans les différents types compositionnels rédigés en tchèque par des locuteurs non natifs (partie 4.1) et d'ensuite identifier et comparer les conjonctions les plus fréquentes dans chacun de ces types (partie 4.2).

##### 4.1. Occurrences des conjonctions dans les différents types de textes

Le tableau 2 présente les nombres de conjonctions dans les différents types compositionnels.

Niveaux du CECR →	A	A	B	B	C	C
Type de textes ↓	Occurrences	i.p.m.	Occurrences	i.p.m.	Occurrences	i.p.m.
<b>Procédé argumentatif</b>	8949	90680,64	8334	90088,53	645	83376,42
<b>Procédé narratif</b>	12075	72287,64	7180	77511,01	951	83751,65
<b>Procédé descriptif</b>	12878	66940,08	5400	66856,51	151	57153,67
<b>Procédé informatif</b>	10290	51424,80	4254	61542,47	251	66648,96

*Tableau 2. Nombres de conjonctions dans les différents types compositionnels du corpus CzeSL-SGT*

Nous avons en outre analysé chacun des types de textes en regard des niveaux du CECR (A-C). Les niveaux A et B présentent les mêmes tendances, à savoir que les plus hautes fréquences de conjonctions apparaissent dans les textes argumentatifs (A : 90 680,64 i.p.m., B : 90 088,53 i.p.m.), puis dans les textes narratifs (A : 72 287,64 i.p.m., B : 77 511,01 i.p.m.), et descriptifs (A : 66 940,08 i.p.m., B : 66 856,51 i.p.m.). Les fréquences les plus faibles sont représentées par les textes informatifs (A : 51 424,80, B : 61 542,47).

Cet ordre correspond à l'interprétation de MISTRÍK (1997), voir partie 1.5. En effet, il considère comme type le plus cohésif les textes explicatifs (à l'intérieur desquels il place les textes argumentatifs), et comme moins cohésif, le type informatif. Il qualifie le type descriptif de non cohésif, mais plus cohésif que le type informatif. En revanche, il considère le type narratif comme cohésif, bien que moins cohésif que l'explicatif (argumentatif).

Dans la catégorie C, le nombre relatif de conjonctions dans les textes était équivalent pour les types argumentatif (83 376,42 i.p.m.) et narratif (83 751,65 i.p.m.), et moindre dans les textes relevant des types informatif (66 648,96 i.p.m.) et descriptif (57 153,67 i.p.m.). Ces résultats peuvent être influencés par le fait que



le corpus CzeSL-SGT présente peu de textes de niveau C. Dans l'ensemble, ces résultats correspondent néanmoins aux hypothèses initiales : les procédés compositionnels narratif et argumentatif sont plutôt cohésifs, tandis que les procédés informatif et descriptif sont plutôt non cohésifs.

Les résultats montrent que le nombre d'occurrences de conjonctions est lié aux caractéristiques d'un procédé compositionnel donné : les types argumentatifs et narratifs apparaissent comme plus complexes du point de vue des connecteurs discursifs. Ces deux types présentent des relations textuelles plus variées et plus complexes que les textes descriptifs et informatifs, qui du point de vue de la construction discursive sont généralement plus simples. En effet, en comparaison avec les types explicatif et narratif, les types descriptif et informatif comprennent moins de relations discursives exprimées explicitement, c'est-à-dire moins de connecteurs.

#### **4.2. Conjonctions utilisées dans les différents types de textes**

Nous avons également analysé les textes relevant des différents types compositionnels en fonction des fréquences des différentes conjonctions. À nouveau nous avons distingué les textes en fonction du niveau des utilisateurs (selon les niveaux du CECR : A,B,C). Voici un récapitulatif des 10 conjonctions (Conj.) les plus fréquentes pour chacun des types compositionnels.

##### **Procédé argumentatif**

<b>A</b>			<b>B</b>			<b>C</b>		
<b>Ordre</b>	<b>Conj.</b>	<b>Occurrences</b>	<b>Ordre</b>	<b>Conj.</b>	<b>Occurrences</b>	<b>Ordre</b>	<b>Conj.</b>	<b>Occurrences</b>
1.	a	3238	1.	a	3051	1.	a	216
2.	že	1544	2.	že	1288	2.	že	82
3.	ale	987	3.	ale	859	3.	ale	66
4.	protože	614	4.	když	567	4.	i	41
5.	když	519	5.	protože	451	5.	když	30
6.	nebo	370	6.	nebo	415	6.	jako	29
7.	proto	279	7.	aby	301	7.	aby	28
8.	i	259	8.	i	289	8.	nebo	28
9.	aby	237	9.	proto	247	9.	proto	14
10.	jako	226	10.	jako	209	10.	než	11

### Procédé narratif

A			B			C		
Ordre	Conj.	Occurrences	Ordre	Conj.	Occurrences	Ordre	Conj.	Occurrences
1.	a	5640	1.	a	3098	1.	a	411
2.	že	1444	2.	že	915	2.	že	152
3.	ale	1348	3.	ale	900	3.	ale	79
4.	protože	924	4.	když	564	4.	když	58
5.	když	793	5.	protože	428	5.	jako	44
6.	proto	432	6.	proto	228	6.	protože	31
7.	nebo	304	7.	jako	195	7.	aby	30
8.	jako	260	8.	nebo	160	8.	proto	22
9.	aby	238	9.	i	152	9.	i	17
10.	i	168	10.	aby	146	10.	než	14

### Procédé descriptif

A			B			C		
Ordre	Conj.	Occurrences	Ordre	Conj.	Occurrences	Ordre	Conj.	Occurrences
1.	a	6942	1.	a	2742	1.	a	79
2.	že	1322	2.	že	532	2.	ale	13
3.	ale	1121	3.	ale	470	3.	když	10
4.	protože	938	4.	protože	383	4.	nebo	8
5.	když	636	5.	když	299	5.	že	7
6.	nebo	489	6.	nebo	237	6.	jako	7
7.	proto	345	7.	jako	156	7.	i	6
8.	jako	316	8.	proto	154	8.	aby	4
9.	aby	161	9.	aby	92	9.	protože	4
10.	než	132	10.	i	77	10.	však	2

### Procédé informatif

A			B			C		
Ordre	Conj.	Occurrences	Ordre	Conj.	Occurrences	Ordre	Conj.	Occurrences
1.	a	5754	1.	a	2098	1.	a	152
2.	že	1086	2.	že	547	2.	ale	24
3.	ale	918	3.	ale	368	3.	že	22
4.	protože	740	4.	protože	283	4.	protože	13
5.	nebo	459	5.	když	227	5.	proto	10
6.	když	391	6.	nebo	200	6.	když	10
7.	proto	265	7.	proto	139	7.	nebo	5
8.	jako	214	8.	aby	84	8.	než	3
9.	i	82	9.	jako	74	9.	kdyby	2
10.	aby	68	10.	i	59	10.	ani	2

Ce panorama montre que les conjonctions les plus fréquentes sont sensiblement les mêmes dans tous les types de textes. Remarquons que, pour les niveaux A et B, les trois conjonctions les plus fréquentes coïncident, quel que soit le type de texte :

la plus fréquente étant *a*, ensuite la conjonction introduisant une proposition complétive *že*, et la troisième *ale*<sup>10</sup>. Au niveau C apparaissent des écarts à ce schéma, ceux-ci pourraient cependant être attribuables au nombre restreint de ces textes dans le corpus.<sup>11</sup> Il n'en demeure pas moins que, quels que soient le niveau du CECR et le procédé compositionnel, la conjonction *a* est la plus fréquente (avec une avance conséquente sur les suivantes).

Cette constatation présente un intérêt, car dans les procédés compositionnels que Mistrík qualifie de cohésifs, c'est-à-dire les procédés narratifs et explicatifs (argumentatifs), on pourrait s'attendre à des relations plus spécifiques entre les propositions (causales par exemples, mais aussi temporelles dans les textes narratifs) que dans les procédés considérés comme non cohésifs, soit dans les textes informatifs descriptifs. L'analyse révèle cependant que les conjonctions les plus fréquentes sont très similaires dans tous les types compositionnels. Parmi les conjonctions ayant un sens plus spécifique (*a* porte un sens très général et présente de larges modalités d'emploi, tandis que nous ne considérons pas *že* comme un connecteur, car il ne définit pas un sens logique spécifique aux relations qu'il établit), c'est la conjonction *ale* qui se trouve la plus fréquente, tous types compositionnels et niveaux confondus. Cette conjonction établit généralement une relation d'opposition. Par ailleurs, la conjonction *protože*, exprimant généralement des relations causales, est très fréquente même dans les types informatifs et descriptifs, pourtant considérés comme non cohésifs. D'un autre côté, la conjonction d'un sens très général *a*, le plus souvent coordonnante, est la plus fréquente dans tous les types compositionnels, y compris dans ceux dits non cohésifs.

Nous avons ensuite comparé les résultats de notre analyse avec ceux d'une étude similaire basée sur les données du Pražský diskurzvní korpus 2.0<sup>12</sup> (PDiT; RYSOVÁ *et al.* 2016). Il s'agit d'un corpus de textes journalistiques en tchèque (rédigés par des locuteurs natifs). Ces données linguistiques ont été analysées du point de vue des outils de cohésion en ciblant les connecteurs discursifs du tchèque. Il ne s'agissait donc pas d'une analyse directe des conjonctions, mais des outils permettant la cohésion en général. Nous pouvons néanmoins comparer au moins partiellement les résultats de cette étude avec les nôtres. L'analyse des données du PDiT a montré (*cf.* RYSOVÁ 2018) que les cinq connecteurs les plus utilisés dans les textes journalistiques sont *a*, *však*, *ale*, *když* et *protože*. Notre analyse des conjonctions dans des textes rédigés par des non-natifs nous amène aux mêmes conclusions.

---

<sup>10</sup> Traduisibles respectivement par « et », « que » et « mais » (N. d. T.).

<sup>11</sup> La catégorie C présente quelques autres différences au niveau des conjonctions présentes. Par exemple la conjonction *i*, ayant généralement une fonction de coordination, présente une faible fréquence dans les différents types compositionnels et les niveaux du CECR, sauf dans les textes explicatifs de niveau C, où elle est la quatrième conjonction la plus fréquente. Les occurrences de *i* représentent dans cette catégorie environ un cinquième des occurrences de la conjonction la plus fréquente, *a*, alors que dans les textes descriptifs de niveau C par exemple, elle n'y représente qu'un treizième. De même, la conjonction *protože*, qui introduit généralement une relation de cause, est fréquente, au niveau C, dans les textes informatifs et narratifs. Toujours à ce même niveau, elle est peu fréquente dans les textes descriptifs et ne fait même pas partie des dix plus fréquentes dans les textes explicatifs.

<sup>12</sup> Qui pourrait être traduit par « Corpus pragois du discours 2.0 » (N. d. T.).

Il est donc manifeste que lors de la rédaction d'un texte, les auteurs utilisent souvent les mêmes outils pour la cohésion : le répertoire des expressions permettant la cohésion dans les textes écrits ne varie que peu, quel que soit le procédé compositionnel et que l'auteur soit natif de tchèque ou non.

### **Conclusion**

Dans cet article, nous nous sommes consacrés à l'étude des modalités d'utilisation des conjonctions dans différents textes. Plus précisément nous nous sommes demandé si des textes relevant de différents genres rédigés en tchèque par des non-natifs diffèrent dans l'utilisation, en nombre et en nature, des conjonctions utilisées comme outils primaires de la cohésion textuelle.

Les résultats de notre étude montrent que la fréquence des occurrences des conjonctions dépend du procédé compositionnel appliqué dans le texte. Les procédés plus complexes (du point de vue de leur construction et de l'organisation du texte), comme les procédés argumentatif et narratif, présentent un plus grand nombre de conjonctions que les types descriptif et informatif.

L'étude montre en outre que les différents procédés compositionnels ne diffèrent que peu dans la nature des conjonctions les plus fréquentes. Quel que soit le procédé compositionnel ou le niveau du CECR, ce sont d'ordinaire les conjonctions *a*, *že*, *ale*, *protože* et *když* qui comptent parmi les cinq conjonctions les plus fréquentes.

En résumé, les différents types compositionnels se distinguent bien par le nombre de conjonctions utilisées (et par les relations qu'elles expriment), mais peu par la nature des conjonctions qui s'y trouvent les plus fréquentes.

L'utilisation dans un texte des conjonctions les plus fréquentes n'est fondamentalement sensible ni au procédé compositionnel ni au niveau linguistique du texte tel que défini par le CECR. Il s'avère que les conjonctions les plus fréquentes sont composées d'un même ensemble (peu variable). Il apparaît dès lors que les auteurs adoptent un répertoire précis et peu étendu de conjonctions qu'ils utilisent intensivement et avec peu de variations.

### **BIBLIOGRAPHIE**

- AIJMER Karin (2002), *English Discourse Particles. Evidence from a corpus*, Amsterdam/Philadelphia, John Benjamins.
- BEČKA Josef Václav (1992), *Česká stylistika*, Prague, Academia.
- Common European Framework of Reference for Languages, CEFR* (2011), Council of Europe, WWW <<https://www.coe.int/en/web/common-european-framework-reference-languages/>>.
- DANEŠ František (1989), Report of R. van de Velde's paper Man, verbal text, inferencing and coherence, in : HEYDRICH Wolfgang, NEUBAUER Fritz, PETOFI János S., SOZER Emil (éds), *Connexity and Coherence, Analysis of Text and Discourse*, Berlin, New York, Walter de Gruyter, p. 228-239.
- FRASER Bruce (1999), What are discourse markers?, *Journal of Pragmatics* 31 (7), p. 931-952.

- GÜNTNER Susanne, KNOBLAUCH Hubert (1995), Culturally Patterned Speaking Practices – The Analysis of Communicative Genres, *Pragmatics* 5, p. 1-32.
- HAJIČ Jan *et al.* (2020), *Prague Dependency Treebank – Consolidated 1.0 (PDT-C 1.0)*. Data/software. Prague, Institute of Formal and Applied Linguistics MFF UK, [www <http://hdl.handle.net/ 11234/1-3185>](http://hdl.handle.net/11234/1-3185).
- HAJIČOVÁ Eva (1993), *Issues of Sentence Structure and Discourse Patterns*, Prague, Charles University.
- HAJIČOVÁ Eva (2012), On Scalarity in Information Structure, *Linguistica Pragensia* 22, p. 60-78.
- KARLÍK Petr (2016), Spojka, in : KARLÍK Petr, NEKULA Marek, PLESKALOVÁ Jana (éds), *Nový encyklopedický slovník češtiny*, Prague, Nakladatelství Lidové noviny, [www <https://www.czechency.org/slovník/SPOJKA>](https://www.czechency.org/slovník/SPOJKA).
- KRČMOVÁ Marie (2016), Slohový postup, in : KARLÍK Petr, NEKULA Marek, PLESKALOVÁ Jana (éd.), *Nový encyklopedický slovník češtiny*, Prague, Nakladatelství Lidové noviny, [www <https://www.czechency.org/slovník/SLOHOVÝ POSTUP>](https://www.czechency.org/slovník/SLOHOVÝ_POSTUP).
- LUCKMANN Thomas (1986), Grundformen der gesellschaftlichen Vermittlung des Wissens: Kommunikative Gattungen, *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 27, p. 191-211.
- MISTRÍK Jozef (1997), *Štylistika*, Bratislava, Slovenské pedagogické nakladateľstvo.
- PRASAD Rashmi *et al.* (2008), The Penn Discourse Treebank 2.0, in : MÖLLER Sebastian *et al.* (éd.), *Proceedings of the 6th International Conference on Language Resources and Evaluation*, Marrakech, Morocco, ELRA, p. 2961-2968.
- PRASAD Rashmi *et al.* (2010), Realization of discourse relations by other means: Alternative lexicalizations, in : HUANG Chu-Ren, JURAFSKY Dan (éds), *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*, Beijing, Tsinghua University Press, p. 1023-1031.
- RICKHEIT Gert, SCHADE Ulrich (2000), Kohärenz und Kohäsion, in : BRINKER Klaus, ANTOS Gerd *et al.* (éds), *Text- und Gesprächslinguistik: Ein internationales Handbuch zeitgenössischer Forschung*, Berlin, Walter de Gruyter, p. 275-283.
- RYSOVÁ Magdaléna (2018), *Diskurzni konektory v češtině: Od centra k periferii*. Prague, Institute of Formal and Applied Linguistics MFF UK, [www <https://ufal.mff.cuni.cz/books/preview/2018-rysova\\_full.pdf>](https://ufal.mff.cuni.cz/books/preview/2018-rysova_full.pdf).
- RYSOVÁ Magdaléna *et al.* (2016), *Prague Discourse Treebank 2.0*. Data/software. Prague, Institute of Formal and Applied Linguistics MFF UK, [www <http://hdl.handle.net/11234/1-1905>](http://hdl.handle.net/11234/1-1905).
- RYSOVÁ Magdaléna, RYSOVÁ Kateřina (2014), The centre and periphery of discourse connectives, in : AROONMANA-KUN Wirote, BOONKWAN Prachya, SUPNITHI Thepchai (éd.), *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computation (PACLIC 2014)*, Bangkok, Department of Linguistics, Faculty of Arts, Chulalongkorn University, p. 452-459.

- RYSOVÁ Magdaléna, RYSOVÁ Kateřina (2018), Primary and secondary discourse connectives: Constraints and preferences, *Journal of Pragmatics* 130, p. 16-32.
- SGALL Petr (1980), Text a reference, *Slovo a slovesnost* 41 (2), p. 140-145.
- SCHIFFRIN Deborah (1987), *Discourse Markers*, Cambridge, Cambridge University Press.
- ŠEBESTA Karel et al. (2014), *AKCES 5 (CzeSL-SGT) Release 2*, LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, www <<http://hdl.handle.net/11234/1-162>>.
- ŠTÍCHA František (2013), *Akademická gramatika spisovné češtiny*, Prague, Academia.
- VIEHWEGER Dieter (1977), Úvahy ke gramatice textu: propoziční pojetí textu proti pojetí orientovanému k jednání, *Slovo a slovesnost* 38 (1), p. 1-8.