

QUELQUES POSSIBILITÉS DE L'ANALYSE DES CONNECTEURS DE CAUSE À L'AIDE D'UNE BASE DE DONNÉES SUPRA CORPORA

ALEXANDER GONCHAROV

Institut des problèmes d'informatique du Centre de recherche fédéral
'Informatique et gestion' de l'Académie des sciences de Russie, Russie

Résumé. L'article présente quelques approches à l'analyse des connecteurs moyennant la base de données supra corpora (BDSC) sur l'exemple du connecteur de cause *potomu čto* (parce que). La BDSC forme une superstructure par rapport au corpus parallèle. Elle permet d'effectuer la recherche et l'annotation des unités linguistiques, de même que l'obtention des statistiques. L'article dresse un tableau succinct des descriptions de relations causales dans diverses typologies. Il démontre que la BDSC contribue à l'élaboration de la liste des relations logico-sémantiques et à la représentation de ces dernières dans le cadre d'un système de classification consistant et applicable au matériau linguistique. Outre cela, les principes de l'annotation utilisés dans la BDSC peuvent mettre en évidence les particularités de l'usage des connecteurs. Les équivalents de traduction de *potomu čto* (litt. parce que) et de *potomu, čto* (litt. parce, que) sont examinés, comparés et classifiés. À la différence de *potomu čto* le connecteur *potomu, čto* établit dans la plupart des cas la relation sur le niveau propositionnel et non pas énonciatif. Ce fait est appuyé par plusieurs différences de traduction telles que (1) la présence de nombreux connecteurs à deux composantes, (2) l'absence de connecteurs exprimant les relations autres que la cause et (3) de l'implication de *potomu, čto* dans la traduction. Les résultats témoignent de l'efficacité des méthodes assurées par la BDSC pour l'étude des connecteurs. Le travail ultérieur permettrait de cataloguer les moyens langagiers exprimant différentes relations logico-sémantiques et de préciser la classification de ces relations.

Mots clés : linguistique contrastive, linguistique de corpus, bases de données supra corpora, connecteurs logiques, relations logico-sémantiques, cause

INTRODUCTION

Considérés comme moyens prototypiques de cohérence textuelle, les connecteurs font l'objet de nombreuses études linguistiques, y compris celles qui sont basées sur des corpus. Le connecteur est une unité lexicale dont la fonction consiste à exprimer une relation logico-sémantique entre deux fragments de texte que cette

unité lie (Inkova-Manzotti, 2001 : 17). Bien que l'étude des connecteurs ait beaucoup changé avec l'apparition des corpus linguistiques et surtout des corpus annotés, la fonctionnalité de ces ressources est souvent limitée aux possibilités de recherche. Pour pouvoir non seulement effectuer la recherche des connecteurs, mais aussi décrire leur fonctionnement, une base de données *supra corpora* (désormais BDSC) a été créée. Le présent article a pour but d'illustrer quelques approches à l'analyse des connecteurs moyennant la BDSC sur l'exemple du connecteur de cause *potomu čto* (parce que). La première partie présentera les outils de l'analyse des connecteurs (la BDSC), la deuxième partie traitera de l'objet de la présente analyse (relations causales et deux connecteurs de cause, *potomu čto* et *potomu, čto*), la troisième partie vise à présenter les équivalents de traduction des connecteurs mentionnés, et enfin la quatrième partie sert à mettre en évidence les particularités du fonctionnement des connecteurs *potomu čto* et *potomu, čto* en russe à travers leurs équivalents.

OUTILS DE L'ANALYSE

La BDSC utilisée pour l'étude des connecteurs contient les textes du sous-corpus parallèle russe-français du Corpus National de la langue russe (NKRYa) d'un volume total de 3,9 millions d'occurrences pour les textes russes et leurs traductions en français et de 1,2 millions d'occurrences pour les textes français et leurs traductions en russe (données au 15.06.2019).

À la différence des corpus classiques, y compris parallèles, la BDSC permet non seulement d'effectuer la recherche des connecteurs, mais aussi de décrire et d'enregistrer les contextes de leurs occurrences sous forme d'annotations bilingues (russe-français et français-russe). C'est la raison pour laquelle les bases de ce type sont appelées *supra corpora*.

Les annotations se présentent sous forme d'un tableau à cinq colonnes dont la première et la troisième contiennent les contextes du connecteur analysé et de son équivalent de traduction, tandis que la deuxième et la quatrième sont réservées aux propriétés du connecteur dans ce contexte et de son équivalent, respectivement. Ces propriétés sont les suivantes : la relation logico-sémantique exprimée par les connecteurs, la structure syntaxique du fragment de texte qu'ils introduisent, leur position dans ce fragment, l'ordre d'enchaînement des fragments de texte, le statut de l'unité lexicale (connecteur, partie d'un connecteur) et la disposition des éléments qui composent le connecteur (à contact ou à distance). Dans la BDSC, les connecteurs sont fixés sous la forme dans laquelle ils sont utilisés dans chaque contexte précis. Cette forme est appelée 'variante discursive' (désormais VD). La VD d'un connecteur et son équivalent forment ce qu'on appelle une correspondance de traduction (désormais CT) dont les propriétés sont annotées dans la dernière colonne (cf. pour les principes d'annotation et la terminologie utilisée Zatsman et al., 2016).

Tableau 1 Exemple d'une annotation

Contexte de la VD dans l'original	VD dans l'original et ses propriétés	Contexte de la VD dans la traduction	VD dans la traduction et ses propriétés	Propriétés de la CT
Petrovič provozilsja za šinel'ju vsego dve nedeli, <i>potomu čto</i> mnogo bylo stegan'ja (Gogol, 1842)	<i>potomu čto</i> <cause propositionnelle> <avec prédication> <initiale> <p CNT q> <CNT> <Contact>	La confection du manteau ne prit que deux petites semaines, et encore <i>parce qu'</i> il devait être ouaté et piqué (Gogol, 1938 ; trad. par Mongault)	<i>parce que</i> <cause propositionnelle> <avec prédication> <initiale> <p CNT q> <CNT> <Sub CNT> <Contact>	<type 1> <traduction congruente>

La BDSC rend également possible l'obtention des données statistiques selon toutes les informations que l'annotation contient (Inkova et Popkova, 2017). Comme l'annotation représente l'unité principale de la BDSC, toutes les données quantitatives incluses dans cet article se rapportent aux annotations de CT et non pas aux occurrences des connecteurs russes. Cette remarque est importante, vu la présence dans le corpus de traductions multiples pour certaines œuvres. En conséquence ce sont les CT (dont plusieurs peuvent avoir le même original et les mêmes propriétés ; alors il s'agit de deux premières colonnes du Tableau 1) qui sont analysées dans notre étude.

OBJET DE L'ANALYSE

L'objet de notre étude effectuée à l'aide de la BDSC est un des connecteurs de cause les plus répandus en russe : *potomu čto* (parce que). Bien que la relation de cause soit étudiée dans nombre de travaux, un traitement différent lui est réservé au sein des approches qui se fondent sur l'utilisation des corpus annotés.

D'après la Théorie de la structure rhétorique (Rhetorical Structure Theory, RST) le groupe de relations causales inclut la cause délibérée et non-délibérée, la conséquence ('result') délibérée et non-délibérée, le but. En même temps, malgré toutes les affinités les relations de démonstration, de justification et de motivation ('evidence', 'justify', 'motivation') sont examinées dans le cadre de deux autres groupes, démonstration et justification ('Evidence and Justify') et facilitation et motivation ('Enablement and Motivation', Mann et Thompson, 1988 : 250). Dans les développements ultérieurs (cf. En ligne 1) la cause délibérée et non-délibérée, la conséquence ('résultat' d'après la terminologie proposée) délibérée et non-délibérée, le but sont inclus dans le groupe de relations dites de contenu tandis que la démonstration, la justification et la motivation sont mentionnées parmi les relations de présentation.

La Théorie des représentations discursives segmentées (Segmented Discourse Representations Theory, SDRT) ne recourt pas aux notions de cause, conséquence et but mais à celles d'explication et conséquence ('explanation' et 'result', Asher et Lascarides, 2003 : 204–206).

Pour remédier à cette hétérogénéité des étiquettes les chercheurs qui développent le corpus annoté à l'Université de Pennsylvanie (Penn Discourse Treebank, PDTB) ont proposé un nouveau standard international de classification des relations discursives appelé ISO DR-Core. Ce standard liste 18 relations, y compris la cause (Bunt et Prasad, 2016). Malgré cette proposition dans un manuel d'annotation ultérieur, ils distinguent déjà trois types de relations causales : cause, cause+croissance et cause+acte de langage ('cause', 'cause+belief' et 'cause+SpeechAct'). Outre cela, pour chaque type sont distingués les contextes exprimant la raison et la conséquence ('reason' et 'result') (Webber et al., 2019 : 19–21).

Cet examen concis et non exhaustif de quelques approches existantes fait voir sur l'exemple de la relation de cause que la tâche de créer un système tant complet et détaillé que libre de contradictions et, par conséquent, permettant d'annoter la structure logique et discursive du texte est loin d'être accomplie. Le chercheur qui essaye d'appliquer l'une de ces typologies se heurte généralement à trois types de difficultés. Tout d'abord, il s'agit de notions trop larges couvrant différents emplois d'unités linguistiques (p. ex. 'cause'), ce qui dissimule leurs différences de fonctionnement. Ensuite, dans les mêmes classifications sont utilisés des termes portant d'un côté sur la structure logique du texte et de l'autre sur la pragmatique du discours (p. ex. 'cause' et 'justify'), ce qui, à la différence du premier cas, masque d'éventuels points communs entre les différents emplois d'un connecteur. Enfin, les critères mis à la base des classifications et permettant d'y inclure certaines et en exclure d'autres, ne sont pas clairs.

Prenant en considération les problèmes énumérés on applique à l'analyse de *potomu čto* la classification des relations logico-sémantiques proposée par Inkova (2017 : 116–120). Selon cette classification, les relations de cause, de conséquence et de but sont trois relations distinctes se basant sur l'opération d'implication. Avec la condition, ces 4 relations forment un groupe de relations correspondant au schéma logique $p \rightarrow q$. En même temps, chaque relation peut en règle générale être établie aux niveaux propositionnel, énonciatif et métatextuel. Outre cela, les principes de la classification permettent de déterminer l'opération de base, la relation logico-sémantique qui l'exploite et le statut des arguments reliés, ce qui décrit simultanément l'organisation logique du texte et sa composante pragmatique sans les mélanger.

Comme dans notre étude il s'agit des VD des connecteurs, il est à noter que *potomu čto* peut être employé de manière autonome ou faire partie des unités complexes à deux (*esli... to potomu čto*, litt. *si... c'est parce que*) ou à plusieurs composantes (*potomu čto... čto... čto... čto*, litt. *parce que... que... que... que*). Les VD de ce type ne seront pas prises en considération ici. En revanche, la différence

entre les VD avec et sans virgule (*potomu, čto*, litt. *parce, que* et *potomu čto*, litt. *parce que*) semble importante. L'article traite donc de deux connecteurs de cause : *potomu, čto* (54 annotations) et *potomu čto* (196 annotations).

Dans la tradition grammaticale russe la différence entre ces deux unités est expliquée en termes de mise en relief et/ou de la structure informationnelle de l'énoncé. Ainsi, Kobozeva (2010 : 112–115) applique à leur analyse la dichotomie 'donné – nouveau' ('given – new') proposée par Chafe (1982 : 281–285). Le donné est défini comme information actualisée lors de l'énonciation et que l'interlocuteur se rappelle. En principe, le même contexte peut admettre l'emploi d'une forme soudée (sans virgule) ou détachée (avec virgule) ; cf. l'exemple de Kobozeva (2010 : 113) :

- a. Vasja ne pošel ↓ v školu (nouveau) ||, *potomu čto* u nego noč'ju podnjalaš' temperatura ↓.
- b. Vasja ne pošel v školu (donné) *potomu* ↑ ||, *čto* u nego noč'ju podnjalaš' temperatura ↓. (Vassia n'est pas allé à l'école *parce qu'il* avait eu de la fièvre dans la nuit.)

C'est la prosodie qui permet de comprendre s'il s'agit d'une information nouvelle ou donnée dans cet exemple. La principale prononcée avec un ton descendant (↓) contient l'information nouvelle (l'interlocuteur n'est pas au courant que Vassia n'est pas allé à l'école) et implique l'emploi d'une forme soudée. La principale prononcée avec un ton ascendant (↑) a le statut de donné (l'interlocuteur est au courant que Vassia n'est pas allé à l'école) et c'est une forme détachée qui le signale à l'écrit. Nous essayerons de préciser cette différence entre les deux formes moyennant une analyse contrastive.

ÉQUIVALENTS DE TRADUCTION

L'utilisation de corpus parallèles permet de décrire la structure du texte original en tenant compte de celle de sa traduction. Pour ce qui est de *potomu* (,) *čto*, notre étude montre qu'il existe au moins 6 groupes d'équivalents de traduction des unités en question :

1. connecteurs de cause ;
2. connecteurs exprimant une relation logico-sémantique autre que la cause ;
3. 'lexicalisations alternatives' ('alternative lexicalizations', terme proposé dans Prasad et al., 2010) ;
4. moyens grammaticaux ;
5. signes de ponctuation ;
6. équivalent zéro (= relation logico-sémantique implicite ou absente à cause des modifications apportées pendant la traduction).

Les équivalents de traduction et leur répartition sont présentés dans les Tableaux 2 et 3. La barre verticale signifie que les éléments constitutifs d'un

connecteur ou d'une composante de connecteur sont disposés à distance mais se trouvent pourtant au sein d'un seul et même fragment de texte. La double barre verticale signifie qu'un connecteur comprend deux ou plusieurs composantes dont chacune introduit un fragment de texte.

Tableau 2 Répartition des équivalents de traduction de *potomu čto* dans la BDSC

Groupe d'équivalents de traduction	Équivalent de traduction	Nombre d'annotations par équivalent	Nombre d'annotations par groupe	% par groupe
Connecteurs de cause	car	65	148	75,6 %
	parce que	63		
	puisque	6		
	faute de	5		
	pour la + <i>adj.</i> raison que	3		
	comme	2		
	mais c'est parce que	1		
	parce que et que	1		
	si c'est que	1		
Autres connecteurs	vu	1	8	4,1 %
	pour	3		
	d'ailleurs	2		
	à moins que	1		
	mais	1		
Lexicalisations alternatives	sinon	1	0	0 %
	---	0		
Moyens grammaticaux	<i>participe passé composé</i>	4	13	6,6 %
	<i>participe présent</i>	4		
	<i>subordonnée</i>	2		
	<i>c'est +pron. qui</i>	1		
	<i>adj. que +être (sous une forme finie)</i>	1		
	<i>participe passé</i>	1		
Signes de ponctuation	<i>deux-points</i>	9	12	6,1 %
	<i>tiret</i>	3		
Équivalent zéro	<i>relation logico-sémantique implicite</i>	13	15	7,6 %
	<i>omission d'un fragment de texte</i>	2		
		196		

Tableau 3 Répartition des équivalents de traduction de *potomu, čto* dans la BDSC

Groupe d'équivalents de traduction	Équivalent de traduction	Nombre d'annotations par équivalent	Nombre d'annotations par groupe	% par groupe
Connecteurs de cause	parce que	35	47	87,3 %
	si c'est que	3		
	parce que et que	2		
	si c'est parce que	2		
	si c'est parce que	2		
	à cause de	1		
	c'est que	1		
	dans la mesure où	1		
Autres connecteurs	---	0	0	0 %
Lexicalisations alternatives	encore faut-il que	1	1	1,8 %
Moyens grammaticaux	<i>subordonnée</i>	2	3	5,5 %
	<i>participe présent</i>	1		
Signes de ponctuation	<i>deux-points</i>	1	1	1,8 %
Équivalent zéro	<i>omission d'un fragment de texte</i>	1	2	3,6 %
	<i>changement de structure d'un fragment de texte</i>	1		
		54		

FONCTIONNEMENT DES CONNECTEURS *POTOMU ČTO* ET *POTOMU, ČTO* EN RUSSE VU À TRAVERS LEURS ÉQUIVALENTS

Comme le fonctionnement des connecteurs russes en question est à observer à travers leurs équivalents français, il faut tout d'abord comparer le contenu des Tableaux 2 et 3. Une des divergences les plus importantes concerne l'emploi de *car* et de *parce que* qui sont les connecteurs de cause les plus utilisés pour traduire *potomu* (,) *čto*. D'une part, le connecteur *car* représente l'équivalent de traduction le plus fréquent pour *potomu čto* et il n'est pas du tout utilisé pour traduire *potomu, čto*. De l'autre part, *parce que* (bien qu'il soit souvent utilisé pour traduire *potomu čto*) sert à traduire la majorité des occurrences de *potomu, čto*.

C'est la raison pour laquelle le fonctionnement de *parce que* et *car* est à examiner afin de le comparer au fonctionnement des VD russes avec et sans virgule. Plusieurs travaux portent sur la comparaison de ces deux connecteurs français. Malgré toutes les réserves les chercheurs du groupe Lambda-1 affirment que *parce que* et *car* ont différentes restrictions syntaxiques (concernant l'ordre

d'enchaînement des fragments de texte, l'emploi dans les questions ou dans les structures enchâssées, la présence d'un quantificateur, etc.) et remplissent différentes fonctions : tandis que *parce que* constitue un opérateur, *car* constitue un marqueur d'actes de paroles (Groupe Lambda-1, 1975). S. Zufferey note que ces connecteurs sont souvent interchangeables, mais qu'il reste toujours une certaine différence d'emploi prototypique entre ceux-ci à l'écrit (Zufferey, 2012). Une étude diachronique de *car* et *parce que* menée par L. Degand et B. Fagard met en parallèle le degré de grammaticalisation d'une unité linguistique et sa capacité d'exprimer les relations subjectives et intersubjectives (Degand et Fagard, 2012).

Dans le cadre de la classification de O. Inkova tous ces faits peuvent être résumés de façon suivante. Le connecteur *parce que* exprime une relation causale établie sur le niveau propositionnel alors que le connecteur *car* exprime une relation causale établie sur le niveau énonciatif. Il est donc logique de supposer que les emplois sur le niveau énonciatif sont caractéristiques plutôt du connecteur *potomu čto* que de *potomu, čto*. L'analyse des contextes confirme cette supposition : on ne trouve que 2 annotations sur 54 (soit 3,6 %) illustrant l'emploi de *potomu, čto* sur le niveau énonciatif alors que pour *potomu, čto* il y en a 65 sur 196 (soit 33,2 %).

Dans ces 2 annotations il s'agit du même contexte en russe (ce n'est que la traduction qui est différente) où la virgule est obligatoire parce que le connecteur est enchâssé dans une structure à énumération. Ce signe de ponctuation reflète donc la prosodie caractéristique du discours direct. Dans cet exemple *potomu, čto* a à chaque fois dans son contexte gauche un fragment qui comporte l'information donnée (elle est exprimée dans la question et n'est même pas répétée dans la réponse), ce qui autorise aussi l'utilisation de la forme détachée.

– Počemu že pis'mo neobhodimo? – sprosil on.

– Počemu? [...] – A *potomu*, [...] *čto* vy ne spali noč', pisali vse dlja menja, ja tože ègoistka! Èto, vo-pervyh... [...] Potom... v pis'me vašem igrajuť mysl', čuvstvo... vy žili ètu noč' i utro ne po-svoemu, a kak hotel, čtoby vy žili, vaš drug i ja, – èto vo-vtoryh, nakonec, v-tret'ih... [...] – V-tret'ih, *potomu, čto* v pis'me ètom, kak v zerkale, vidna vašà nežnost', vašà ostorožnost', zabota obo mne [...].

(Gontcharov, 1848–1859)

– Pourquoi la lettre était indispensable? demanda-t-il.

– Pourquoi? [...] *Parce que*, [...] premièrement vous n'avez pas dormi de la nuit, vous avez écrit pour moi; moi aussi je suis égoïste. [...] Deuxièmement... Dans votre lettre on lit une pensée, des sentiments... Vous avez vécu cette nuit et cette matinée non comme d'habitude, mais comme votre ami et moi nous voulions vous voir vivre. Enfin, troisièmement... [...] – Troisièmement, *parce que* cette lettre reflète, comme un miroir, votre tendresse, votre prudence, votre souci pour moi [...].

(Gontcharov, 1988 ; trad. par Jurgenson)

Les connecteurs composés sont beaucoup plus utilisés pour traduire *potomu, čto* (16,7 % contre 1 % pour *potomu čto*). Ceux-ci rendent possible la mise en relief de la principale à l'aide de *si... c'est* et sont ainsi plus proches des formes détachées en russe utilisées également avec cette fonction.

Otčego že ne potrudit'sja? – skazal on, vyražaja *ètim to, čto* on ne tjagotitsja svoim trudom imenno *potomu, čto* neset ego dlja umirajušego *človeka* [...]. (Tolstoï, 1882–1886)
 'Alors pourquoi je plaindrais ma peine?' voulant dire par là que s'il ne plaignait pas sa peine, *c'était parce qu'il* la dépensait pour un mourant [...]. (Tolstoï, 1997 ; trad. par Flamant)

Les connecteurs exprimant une relation logico-sémantique autre que la cause ne sont utilisés que pour traduire *potomu čto*. Il s'agit des relations de but (*pour*), 'contre attente' (*d'ailleurs, mais*), réserve (*à moins que*), alternative négative (*sinon*). La relation de cause dans ces cas-là est affaiblie, comme le montre l'exemple qui suit où la sémantique causale est remplacée par celle de 'contre attente' (dans la traduction de L. Jurgenson (1988) dans le même contexte *potomu čto* est traduit par deux-points).

Zahar! – dolgo kričal on, *potomu čto* Zahar sidel s kučerami u vorot, obrašennyh v pereulok. (Gontcharov, 1848–1859)
 Zakhar! Il appela longtemps son valet, *mais* celui-ci était assis avec des cochers près de la porte cochère. (Gontcharov, 1959 ; trad. par Adamov)

En ce qui concerne l'équivalent zéro, pour *potomu čto* c'est le troisième équivalent de traduction d'après sa fréquence (bien qu'il soit beaucoup moins utilisé que *car* et *parce que*). Il s'agit le plus souvent de cas où la relation de cause est implicite mais le connecteur aurait pu être ajouté.

Kupili sukna očen' horošego – i ne mudreno, *potomu čto* ob ètom dumali eše za polgoda prežde i redkij mesjac ne zahodili v lavki primenjat'sja k cenam [...]. (Gogol, 1842)
 Ils y achetèrent de l'excellent drap, cela se conçoit ; on ne cessait d'y penser depuis six mois et on ne laissait pas passer un mois sans entrer dans les magasins pour s'informer des prix. (Gogol, 1925 ; trad. par de Schloezer)

Pour *potomu, čto* on n'a que deux exemples d'équivalent zéro. Dans les deux cas il ne s'agit pas d'une relation implicite, mais d'une omission du fragment correspondant dans la traduction ou de sa complète reformulation.

CONCLUSION

La BDSC fournit des outils susceptibles de préciser la sémantique, le fonctionnement et l'inventaire des équivalents de traduction des connecteurs, ce qui a été montré sur l'exemple de *potomu (,) čto*. L'analyse rend évidente une tendance vers

la corrélation entre trois paires de paramètres : le donné et le nouveau, les emplois sur les niveaux propositionnel et énonciatif, et finalement les formes détachée et soudée du connecteur. Cela dit, grâce à sa fonction classificatoire la BDSC donne la possibilité non seulement d'étudier telle ou telle unité linguistique mais d'unifier les approches existantes dans le cadre d'un seul système. Enfin, l'étude outillée des connecteurs dans le cadre de ce système permet de préciser la classification des relations logico-sémantiques et de cataloguer les moyens de les exprimer dans les langues mises en comparaison. L'analyse dont les principes ont été tracés dans le présent article ne prétend pas être exhaustive, elle peut être précisée et continuée sur l'exemple d'autres connecteurs et d'autres relations logico-sémantiques.

BIBLIOGRAPHIE

- Asher, N., Lascarides, A. (2003) *Logics of Conversation*. Cambridge: Cambridge University Press.
- Bunt, H., Prasad, R. (2016) *ISO DR-Core (ISO 24617-8): Core Concepts for the Annotation of Discourse Relations*. Available from <https://let.uvt.nl/general/people/bunt/docs/ISA-12-DR-core-revised.pdf> [Accessed on 10 June 2019].
- Chafe, W. (1982) 'Dannoe, kontrastivnost', opredelennost', podležašee, topiki i točka zrenija [Donné, contrastivité, définitude, sujet, topicalité et point de vue]. In A. E. Kibrik (ed.) *Novoe v zarubežnoj lingvistike, XI* (pp. 277–316). Moscow: Progress.
- Degand, L., Fagard, B. (2012) Competing connectives in the causal domain French *car* and *parce que*. *Journal of Pragmatics*, 44: 154–168.
- Groupe Lambda-1 (1975) *Car, parce que, puisque*. *Revue Romane*, 10: 248–280.
- Inkova, O. (2017) Le relazioni logico-semantiche tra gli enunciati: una proposta di classificazione. In M. di Filippo, F. Esvan (eds.) *Studi di linguistica slava* (pp. 105–123). Napoli: Il Torcoliere.
- Inkova, O., Popkova, N. (2017) Statistical Data as Information Source for Linguistic Analysis of Russian Connectors. *Informatics and Applications*, 11(3): 123–131.
- Inkova-Manzotti, O. (2001) *Konnektory protivopostavlenija vo francuzskom i ruskom jazykah. Sopostavitel'noe issledovanie [Connecteurs d'opposition en français et en russe: une étude contrastive]*. Moscow: Informelektro.
- Kobozeva, I. (2010) Parametry var'irovanija pokazatelej mežfrazovyh semantičeskikh svjazej v ruskom tekste [Variation des marqueurs de liens sémantiques interphrastiques dans le texte russe et ses paramètres]. *L'analisi linguistica e letteraria XVIII*, 107–122.
- Mann, W. C., Thompson, S. A. (1988) Rhetorical structure theory: Towards a functional theory of text organization. *Text & Talk*, 8(3): 243–281.
- Prasad, R., Joshi, A., Webber, B. (2010) Realization of discourse relation by other means: Alternative lexicalizations. *Proceedings of COLING 2010, 23rd International Conference on Computational Linguistics, Posters Volume, 23–27 August 2010* (pp. 1023–1031).
- Webber, B., Prasad, R., Lee, A., Joshi, A. (2019) *The Penn Discourse Treebank 3.0 Annotation Manual*. Available from <https://catalog.ldc.upenn.edu/docs/LDC2019T05/PDTB3-Annotation-Manual.pdf> [Accessed on 10 June 2019].
- Zatsman, I., Inkova, O., Kruzhkov, M., Popkova, N. (2016) Predstavlenie krossjazykovykh znaniy o konnektorah v nadkorpusnyh bazah dannyh [Représentation de connaissances cross-linguistiques sur les connecteurs dans les bases de données supra corpora]. *Informatics and Applications*, 10(1): 106–118.
- Zufferey, S. (2012) 'Car, parce que, puisque' revisited: Three empirical studies on French causal connectives. *Journal of Pragmatics*, 44: 138–153.

SITOGRAPHIE

[En ligne 1] Disponible sur <http://www.sfu.ca/rst/07french/definitions.html> [consulté le 15 juin 2019].

SOURCES ANALYSÉES

- Gogol, N. (1842) *Šinel'* [*Le Manteau*].
 Gogol, N. (1925) *Le Manteau*. Trad. par B. de Schloezer.
 Gogol, N. (1938) *Le Manteau*. Trad. par H. Mongault.
 Gontcharov, I. (1848–1859) *Oblomov*.
 Gontcharov, I. (1959) *Oblomov*. Trad. par A. Adamov.
 Gontcharov, I. (1988) *Oblomov*. Trad. par L. Jurgenson.
 Tolstoï, L. (1882–1886) *Smert' Ivana Il'iča* [*La Mort d'Ivan Ilitch*].
 Tolstoï, L. (1997) *La Mort d'Ivan Ilitch*. Trad. par F. Flamant.

APPROACHES TO ANALYSIS OF CAUSAL CONNECTIVES WITH THE HELP OF A SUPRACORPORA DATABASE

Abstract. The paper presents a number of approaches to analysis of connectives with the help of a supracorpora database (SCDB) using the Russian causal connective *potomu čto* (because) as an example. The SCDB acts as a corpus-based superstructure that makes it possible to find and annotate various linguistic units in a parallel corpus and to obtain statistics. The paper provides a brief overview of how the causal relation is described in existing typologies. It demonstrates that the SCDB helps to refine the list of logical-semantic relations and to represent these relations as components of a single consistent classification system based on language material. Furthermore, the annotation principles and the classification system of the SCDB make it possible to uncover specific features that connectives display in their usage. Translation equivalents for the connectives *potomu čto* (lit. because that) and *potomu, čto* (lit. because, that) are examined, compared and classified. As opposed to *potomu čto*, the connective *potomu, čto* in the majority of cases establishes a causal relation between propositions and is rarely used in the illocutive sense. This observation is supported by differences in translation, such as the presence of numerous two-component connectives as well as the absence of connectives expressing non-causal logical-semantic relations and the absence of implicit causal relation. These results demonstrate that the SCDB-based methods of investigating connectives are rather promising. Further work in this direction will be aimed at creating a catalogue of language means capable of expressing different relations and at refining the classification of logical-semantic relations.

Key words: contrastive linguistics, corpus linguistics, supracorpora databases, logical connectives, logical-semantic relations, cause

Alexander Goncharov (chercheur junior) travaille actuellement au Centre de recherche fédéral 'Informatique et gestion' de l'Académie des sciences de Russie. Ses intérêts de recherche portent sur la linguistique de corpus, la linguistique contrastive, l'organisation logique du texte, les relations logico-sémantiques implicites, les classifications des relations intratextuelles en général. E-mail : a.gonch48@gmail.com