

C'est la définition de quel mot?

Tester la validité des définitions lexicographiques pour un dictionnaire d'apprentissage

Jasmina Milićević

Dalhousie University & OLST – Université de Montréal

jmilicev@dal.ca

Résumé

Cet article décrit un test effectué dans le but de vérifier la validité d'un sous-ensemble de définitions lexicographiques élaborées pour un dictionnaire d'apprentissage du français langue seconde. Le test consistait à soumettre aux participants les définitions d'une vingtaine de lexies françaises dénotant des sentiments et à leur demander d'identifier les lexies correspondantes. Les résultats du test étaient censés mettre en évidence des lacunes des définitions proposées et – à travers une analyse des difficultés éprouvées par les participants – des pistes à suivre afin de les améliorer. Ces résultats nous ont permis, notamment, de faire une description plus fine des différences entre chaque lexie testée et des lexies sémantiquement apparentées avec lesquelles l'utilisateur du dictionnaire risque de confondre celle-ci.

1 L'objectif et la méthodologie du test

Le présent article décrit un test effectué dans le but de vérifier la validité d'un sous-ensemble de définitions lexicographiques élaborées dans le cadre du projet de dictionnaire *Dire autrement* (Milićević & Hamel 2007, Hamel & Milićević 2007), un dictionnaire électronique d'apprentissage du français langue seconde de niveau intermédiaire-avancé. *Dire autrement* est un dictionnaire de type explicatif et combinatoire « didactisé », c'est-à-dire 1) élaboré selon la méthodologie de la *Lexicologie explicative et combinatoire*, ou la LEC (Mel'čuk *et al.* 1995, Mel'čuk 2006), et 2) mis sous une forme plus facilement accessible pour les apprenants (sur les définitions pédagogiques à la LEC voir, notamment, Milićević 2008 et Apresjan *et al.* à paraître ; sur l'encodage convivial des collocations, voir Popovic 2003).

Les définitions de la LEC sont des définitions analytiques, ou paraphrastiques, c'est-à-dire faites par la décomposition du sens de la lexie décrite L en termes de sens plus simples que le sens de L. (Le sens 's₁' est plus simple que le sens 's₂' si 's₁' figure dans la décomposition de 's₂' et que l'inverse n'est pas vrai.)

À titre d'exemple, les définitions du verbe SE TAIRE (*Jean se taisait*) et du nom SECRET (*Je n'ai pas de secrets devant toi*) se présentent comme suit.

DEFINI	DEFINISSANT
<i>X se tait</i>	= 'individu X, qui est censé parler, ne parle pas'
<i>secret de X concernant Z(Y)</i>	= 'information qu'a individu X concernant fait Z lié à individu Y telle que X ne doit pas la communiquer aux autres'

Dans le cas d'une lexie prédicative, le défini est la *forme propositionnelle*, une expression contenant la lexie L et ses actants sémantiques (représentées par des variables X, Y, Z, etc.), et le définissant est la décomposition sémantique proprement dite du sens de L. Les composantes sémantiques imprimées en police spéciale dans les définitions ci-dessus (individu, information et fait) sont des caractérisateurs taxinomiques, ou étiquettes sémantiques (sur leur utilisations dans le cadre de la LEC, voir Milićević, 1997 et Polguère, 2004).

Pour vérifier la validité de nos définitions, nous avons eu recours au test de reconnaissance, un test dans lequel il s'agit d'identifier la lexie L en lisant la définition de L. Autrement dit, il faut trouver le mot

à partir de la description de son sens. L'hypothèse à la base du test était qu'une définition adéquate d'une lexie L devrait permettre au locuteur d'identifier L. Une définition adéquate de L est telle que chaque composante en est nécessaire et l'ensemble de composantes est suffisant pour couvrir tous les emplois de L ; c'est ce qu'on connaît sous le nom de *principe d'adéquation*.

Les définitions dans un dictionnaire d'apprentissage doivent en outre être conformes au principe de convivialité. Ceci veut dire que leur format – la structure et le langage des définitions – doit être adapté aux besoins d'apprenants (*grosso modo*, leur niveau d'apprentissage). Nos définitions sont légèrement didactisées par rapport aux définitions standard de la LEC, notamment celles qu'on trouve dans le *Dictionnaire explicatif et combinatoire du français contemporain*, ou DECFC (Mel'čuk *et al.* 1984-1988-1992-1999) ; une comparaison des deux formats de définition sera donnée plus loin. Notre test ne tenait pas compte de l'aspect convivialité des définitions (il n'était pas conçu pour cela). Il est clair, cependant, que la « lisibilité » des définitions dans un dictionnaire d'apprentissage doit être prise en compte lorsqu'on juge de leur adéquation.

Insistons que l'identification réussie n'est pas LA preuve de l'adéquation : il ne s'agit que d'une indication parmi d'autres. (Ainsi, le test de reconnaissance ne permet d'évaluer qu'un aspect de la validité des définitions ; voir immédiatement ci-dessous). Pour illustrer ce fait, considérons les deux définitions suivantes.

RANCUNE (<i>Nouveau Petit Robert</i>)	Souvenir tenace qu'on garde d'une offense, d'un préjudice, avec de l'hostilité et le désir de vengeance.
---	--

Cette définition permet d'identifier assez facilement la lexie visée ; pourtant, son adéquation peut être remise en question. Notamment, la composante 'souvenir', qui est une instance de 'phénomène psychologique', n'est pas le genre prochain de RANCUNE, qui, elle, est une instance de 'sentiment'. On peut également se demander si les composantes 'vengeance' et 'hostilité' sont nécessaires. L'acceptabilité de la phrase *Je lui garde la rancune pour ce qu'il m'a fait, mais je ne cherche pas de vengeance* indique que la rancune ne pousse pas nécessairement à la vengeance (et que, donc, la composante correspondante ne doit pas figurer dans la définition de cette lexie). La situation est moins claire avec la composante 'hostilité' (si l'expression '*rancune sans hostilité*' a l'air bizarre, la disjonction *rancune ou hostilité* semble normale), mais nous sommes plutôt d'avis que la rancune ne se manifeste pas nécessairement dans le comportement de la personne qui la ressent. Garder la rancune à quelqu'un c'est lui en vouloir de façon permanente pour ce qu'il nous a fait et vouloir qu'il lui arrive quelque chose de semblable (cf. la locution GARDER UNE DENT, qui est la verbalisation exacte de RANCUNE) ; pour notre définition de RANCUNE, voir plus loin, la sous-section 3.1.

AFRAID (Wierzbicka 1992 : 178)

X thinks something like this:
Something bad can happen.
I do not want this.
I want to do something because of this.
I do not know what I can do.
Because of this, X feels something bad.

Cette définition présente le cas inverse : la reconnaissance de la lexie correspondante est moins facile (s'agit-il de la lexie AFRAID, WORRIED ou APPREHENSIVE, par exemple ?), mais son adéquation ne peut pas être remise en question aussi facilement que celle de la définition du *Petit Robert* ci-dessous.

La reconnaissance de la lexie par la lecture de sa définition (à l'aide de l'intuition linguistique, des connaissances sur la langue¹ et de la logique générale) ne prouve donc pas que la définition est bonne : elle démontre seulement que la définition est suffisante pour établir les distinctions nécessaires (entre L et les lexies apparentés à L). Il s'agit, plus précisément, des deux types suivants de distinctions :

¹ Nous pensons ici aux connaissances spécialisées sur la langue, c'est-à-dire les connaissances en linguistique, que nous opposons à la connaissance (= maîtrise) de la langue, cette dernière ne présupposant pas les connaissances en linguistique.

- Distinctions entre la lexie L et ses quasi-synonymes.

Les quasi-synonymes sont des lexies qui partagent des composantes sémantiques importantes et sont substituables dans au moins quelques contextes ; par exemple, RESPECT#1 [$\approx$ 'X croit que Z(Y) a une grande valeur sociale ou morale et que, à cause de cela, X doit prendre en considération les opinions de Y'] et ADMIRATION [$\approx$ 'X respecte Z(Y) et voudrait être comme Y']. Cf. aussi les séries suivantes de quasi-synonymes : MÉPRIS ~ DÉDAIN ~ IRRESPECT ; MÉCONTENTEMENT ~ INSATISFACTION ~ DÉPLAISIR ; HONTE ~ DÉSHONNEUR ~ HUMILIATION.

- Distinctions entre L et ses L proches.

Ce sont des lexies qui, sans être des quasi-synonymes (elles ne sont pas substituables en contexte), sont sémantiquement apparentées. Comme un exemple de lexies de sens proche, on peut citer MÉPRIS ($\approx$ 'X croit que (Z)Y n'a pas de valeur sociale ou morale et se croit meilleur que Y' ; Y est un individu) et DÉGOÛT ($\approx$ 'X perçoit Y comme très déplaisant et veut l'éviter' ; Y peut être n'importe quoi) ; il y a dans le mépris quelque chose de physiquement désagréable pour X (quelque chose qui se rapproche du dégoût), mais cette parenté est à notre avis trop lointaine et ne devrait pas être reflétée dans les définitions des lexies correspondantes. Cf. aussi les séries suivantes : MÉCONTENTEMENT ~ DÉCEPTION ~ FRUSTRATION ; HONTE ~ GÊNE ; PEINE ~ DÉPRIME ~ GÊNE ~ MALAISE.

Nous nous attendions à ce que notre test mette en évidence des lacunes des définitions proposées et – à travers une analyse des difficultés éprouvées par les participants dans la reconnaissance des lexies correspondantes – des pistes à suivre afin de les améliorer. Notamment, la confusion entre L et un de ses quasi-synonymes indiquerait des problèmes (relativement) mineurs avec la définition de L, qu'il s'agirait alors de peaufiner ; la confusion entre L et un de ses sens proches signalerait des problèmes plus graves avec la définition de L et la nécessité d'une correction plus sérieuse de cette dernière.

Nous voulions également tester l'impact des connaissances en linguistique sur la reconnaissance des lexies, l'hypothèse étant que les connaissances en linguistique (plus précisément, les connaissances en sémantique/lexicologie et tout particulièrement la familiarité avec le formalisme d'écriture des définitions) facilitent la reconnaissance.

2 Le contenu du test et les participants

Le test consistait à identifier, en lisant leurs définitions respectives, les lexies suivantes :

- 17 lexies du champ sémantique 'sentiments':²

COLÈRE#1, HAINE#1.a (envers qqn), HAINE#1.b (envers qqch), HONTE, HOSTILITÉ#1, INDIGNATION, IRRITATION, MÉCONTENTEMENT, MÉPRIS#1 (pour qqn), PEINE, PEUR, RANCUNE, REGRET#1.1 (de qqch), REGRET#1.2 (d'avoir fait qqch), REMORDS, REPENTIR, RESPECT#1 (pour qqn).

- 1 lexie « hors champ » :

HOSTILITÉ#2 **pl. tant** 'combats armés ...', liée par polysémie à HOSTILITÉ#1 'sentiment négatif...'.
Ces lexies appartiennent toutes au français courant (même si elles n'ont pas la même fréquence dans les textes) et sont censées être connues de tout locuteur adulte du français.

Les lexies dénotant les sentiments ont jusqu'à trois actants sémantiques :

- 1) l'individu X qui éprouve le sentiment;
- 2) la situation Y (possiblement causée par l'individu X lui-même ou par l'individu Z) que X évalue comme bonne/mauvaise, désirable/indésirable, etc., pour X, cette évaluation causant le sentiment de X;
- 3) l'individu Z qui est la « source » et/ou la « cible » du sentiment de X.

² Plus précisément, le champ sémantique en question est constitué des lexies dénotant les sentiments (AMOUR, HAINE, JOIE, TRISTESSE), les émotions (COUP DE FOUDRE, ÉMOI) et les attitudes émotionnelles (RESPECT, ADMIRATION). Il s'agit d'un champs vaste, avec de multiples liens entre lexies et un taux de polysémie élevé – donc, difficile du point de vue de l'acquisition et de la description lexicographique. Voici, à titre d'illustration, trois schémas de polysémie fréquents dans le champ en cause : « sentiment ~ objet du sentiment » (JOIE, ADMIRATION), « sentiment ~ manifestation/instance du sentiment » (COLÈRE, PEUR) et « sentiment envers qqn ~ sentiment envers qqch » (AMOUR, HAINE).

Les définitions des lexies dénotant les sentiments sont élaborées selon le même schéma et contiennent ce qu'on appelle des *blocs de définition standard* (voir Iordanskaja & Mel'čuk 1990, Apresjan & Apresjan 1995, ainsi que Wierzbicka 1992 et 1999) :

- 1) la caractérisation du sentiment (plaisant, déplaisant, fort, etc.) ;
- 2) l'évaluation, par X, de la situation Y/son participant Z qui cause le sentiment ;
- 3) optionnellement, la réaction de X au sentiment ;
- 4) la composante « tel que », qui indique la nature commune du sentiment décrit (cette composante est nécessaire parce que nous ne pouvons pas appréhender les sentiments des autres que par une comparaison avec ce que nous ressentons dans des situations semblables).

Voici, à titre d'illustration, la définition de la lexie PEUR (*Tous les enfants ont peur de l'obscurité, du loup, des monstres, des voleurs*) :

FORME PROPOSITIONNELLE	<u>~ de l'individu X DE Y</u>
CARACTERISATION DU SENTIMENT	'Sentiment négatif de X
EVALUATION DE LA SITUATION (QUI CAUSE LE SENTIMENT)	causé par le fait suivant : X perçoit ou imagine Y comme dangereux pour X ; ceci cause chez X une forte envie d'éviter Y.
REACTION AU SENTIMENT	Si ce sentiment de X est suffisamment intense, il peut causer que X perde le contrôle de lui-même.
COMPOSANTE « TEL QUE »	Ce sentiment est le sentiment que les gens ont normalement dans des situations semblables.'

La plupart des définitions utilisées pour le test ont été reprises du DECFC. Dans beaucoup de cas, la définition reprise a été remaniée (= le contenu en a été modifié) indépendamment de la mise en forme « apprenant ». A titre de comparaison, la définition de PEUR dans le DECFC II: 276 se présente comme suit :

Peur de X de Y ≡ 'Émotion désagréable de X causée par le fait suivant : X croit que l'événement (lié à) Y (concernant l'être Y précieux pour X) qui lui est indésirable est très probable et que X n'est pas capable de s'opposer à Y, et X veut échapper à Y; cette émotion est telle qu'en augmentant, elle cause que X perde la maîtrise de lui-même; elle est celle qu'on a normalement dans de pareilles situations.'³

Neuf personnes ont participé au test, dont huit francophones (1 Québécoise, 3 Français, 3 Acadiens, 1 Camerounais) et un russophone ayant une excellente maîtrise du français. Ces personnes comptaient 3 linguistes-professeurs d'université, 1 doctorant en linguistique, 2 doctorants en études françaises-filière linguistique, 1 étudiant du premier cycle en études françaises-filière linguistique, et 2 personnes sans formation en linguistique. Il s'agissait donc de cinq variétés du français, dont une non native, et de niveaux de connaissances en linguistique très différents ; en outre, on avait affaire à des niveaux de connaissance DE la langue différents (cette connaissance étant plus difficile à évaluer/quantifier dans un test comme le nôtre).

3 Les résultats du test

3.1 Les résultats par lexies

Dans cette sous section, nous indiquons les informations suivantes : le sommaire des réponses (Tableau 1), l'ensemble des réponses (Tableau 2), la distribution des réponses par lexie (Tableau 3), les quasi-synonymes et les sens proches des lexies visées trouvés dans les réponses (Tableau 4).

³ Nous traitons la lexie PEUR comme une instance de sentiment – plutôt que d'émotion – en prenant en compte les données sur la fréquence des collocations *sentiment de peur* (44. 400 d'occurrences sur Google) et *émotion de peur* (4. 200 d'occurrences).

Aucune réponse	10
Réponses correctes	71
Réponses incorrectes	81
• Quasi-synonymes de L	24
• Sens proches de L	20
• “Off mark”	37
Total des réponses	152

Tableau 1 : Sommaire des réponses

Comme on peut le constater, 81 réponses, soit plus que la moitié, sont incorrectes. Parmi celles-ci, plus que la moitié sont des quasi-synonymes et des sens proches des lexies visées, distribuées de façon (presque) égale – 24 et 20, respectivement. Les 37 réponses complètement ratées (= « off mark ») sont de deux variétés.

Les réponses « off-mark » sont pour la plupart constitués des lexies sémantiquement plus éloignées de la lexie visée L que ses quasi-synonymes et sens proches, mais qui

1) partagent avec L certaines composantes, possiblement dans une position différente dans la définition (par ex., HOSTILITÉ#1 [‘désir de faire mal à Y’] et HAINE#1.a [‘désir que qqch de mauvais arrive à Y’]; HONTE [‘sentiment ... ‘X considère Y(X) comme inadéquat, ce qui le pousse à **cacher** l’existence de Y’] ~ SECRET [‘information sur Y que X est censé **cacher** des autres’]) ou

2) sont pragmatiquement reliées à L (par ex., on peut rapprocher RESPECT#1 et AMOUR grâce au fait que les deux sentiments, sans s’impliquer mutuellement, « vont ensemble » dans la réalité ; REGRET et NOSTALGIE se rapprochent par le fait d’être orientés vers des expériences plaisantes révolues, quoique de nature différente).

Un petit nombre de réponses « off mark » est constitué de lexies dont le sens n’a aucun lien avec le sens de la lexie visée (par ex., HAINE#1.a et DÉCEPTION).

	1	2	3	4	5	6	7	8	9	
COLÈRE#1	✓	haine	✓	ressentiment	vengeance	vengeance	vengeance	✓	✓	4/9
HAINE#1.a	mépris	?	déception	✓	malveillance	dégoût	✓	✓	✓	4/9
HAINE#1.b	✓	?	✓	désapprobation	intolérance	irritation	désespoir	aversion	✓	3/9
HONTE	✓	✓	impudeur	gêne	✓	secret	✓	✓	✓	6/9
HOSTILITÉ#1	animosité	✓	indignation	irritation	antipathie	haine	désaccord	antipathie	✓	2/9
HOSTILITÉ#2	✓	guerre	guerre	guerre	différend	guerre	guerre	guerre	✓	2/9
INDIGNATION	dégoût	?	réprobation	révolte	regret	crime	justice	dégoût	✓	1/9
IRRITATION	✓	✓	✓	enervement	colère	?	angoisse	anxiété	✓	4/9
MÉCONTENTEMENT	✓	déception	déception	déception	souffrance	rancune	Insatisfaction	frustration	✓	2/9
MÉPRIS#1	indignation	dégoût	✓	✓	dédain	jalousie	✓	désapproba-	✓	4/9
PEINE	désespoir	✓	mal	déprime	déprime	?	gêne	malaise	?	1/9
PEUR	✓	✓	✓	✓	✓	?	✓	✓	✓	8/9
RANCUNE	✓	✓	✓	✓	vengeance	✓	✓	✓	✓	8/9
REGRET#1.1	peine	nostalgie	nostalgie	nostalgie	frustration	amour	✓	✓	✓	3/9
REGRET#1.2	✓	✓	✓	✓	?	revanche	✓	✓	✓	7/9
REMORDS	✓	regret	✓	culpabilité	dépréciation	regret	✓	regret	✓	4/9
REPENTIR	✓	remords	✓	?	✓	pardon	résolution	?	✓	4/9
RESPECT#1	✓	✓	admiration	admiration	admiration	amour	✓	approbation	✓	4/9
	12/18	8/18	9/18	5/18	3/18	1/18	9/18	7/18	17/18	

Tableau 2 : Ensemble des réponses

Pour ce qui est de la distribution des réponses par lexie, les meilleurs résultats reviennent aux lexies PEUR et RANCUNE : 8 réponses correctes sur 9 dans les deux cas. REGRET#1.2 et HONTE ont également eu un taux de reconnaissance acceptable : 7 et 6, respectivement. Les pires résultats reviennent aux lexies INDIGNATION et PEINE : 1 réponse correcte sur 9 dans les deux cas.

	Incorrectes			Correctes	Aucune
	QSyn	SProche	“Off-mark”		
PEUR	0	0	0	8	1
RANCUNE	0	0	1	8	0
REGRET#1.2	0	0	1	7	1
HONTE	1	0	2	6	0
REMORDS	4	0	1	4	0
RESPECT#1	3	0	2	4	0
IRRITATION	2	0	2	4	1
MÉPRIS	1	1	3	4	0
REPENTIR	1	0	2	4	2
HAINE#1.a	0	1	3	4	1
COLÈRE#1	0	2	3	4	0
HAINE#1.b	1	1	3	3	1
REGRET#1.1	0	3	3	3	0
HOSTILITÉ#2	6	1	0	2	0
HOSTILITÉ#1	3	1	3	2	0
MÉCONTENTEMENT	1	4	2	2	0
PEINE	1	4	1	1	2
INDIGNATION	0	2	5	1	1
	24	20	37	71	10

Tableau 3 : Distribution des réponses par lexie

Voici les définitions de RANCUNE et d’INDIGNATION telles que présentées aux participants du test (voir aussi la définition de PEUR, donnée précédemment). La version corrigée de la définition d’INDIGNATION peut être trouvée à la section 4.

RANCUNE (*Elle continuait à nourrir contre Florent une rancune terrible*)

~ DE L’individu X ENVERS L’individu Y À CAUSE DU fait Z(Y) ≡

‘Sentiment négatif de X envers Y

causé par le fait suivant:

Y a fait à X un Z relativement mauvais

dont X se souvient encore et que X ne pardonne pas à Y;

ceci cause chez X le désir

que quelque chose similaire à Z arrive à Y.

Ce sentiment est le sentiment que les gens ont dans des situations semblables.’

INDIGNATION (*mon indignation contre ces crimes <ces criminels>*)

~ DE LA personne X CONTRE acte/comportement Z DE LA personne Y ≡

‘Très fort sentiment négatif de X envers Z de Y

causé par le fait suivant:

X considère Z de Y

comme moralement ou socialement inacceptable;

ceci cause chez X le désir de faire quelque chose

pour que les Z semblables cessent d’avoir lieu.

Ce sentiment est le sentiment que les gens ont normalement dans des situations semblables.’⁴

⁴ L’étiquette sémantique personne est une étiquette bien distincte de l’étiquette individu : elle désigne un ensemble d’individus ayant une fonction sociale ou politique (SYNDICAT, GOUVERNEMENT, etc.) ; cf. l’expression *personne morale*.

	QSyn	SProche
COLÈRE#1	emportement ; “fam” <i>rogne</i> ; <i>rage</i> ; <i>fureur</i>	<i>ressentiment</i> ; <i>haine</i>
HAINE#1.a	<i>hostilité#1, animosité</i>	<i>malveillance</i>
HAINE#1.b	aversion ; “lit” <i>exécration</i> ; <i>dégoût</i>	<i>intolérance</i>
HONTE	<i>déshonneur</i> ; <i>humiliation</i>	<i>gêne</i>
HOSTILITÉ#1	animosité ; anthipatie	<i>haine</i>
HOSTILITÉ#2	guerre	<i>différend</i>
INDIGNATION	<i>scandale</i> ⁵	<i>réprobation</i> ; <i>révolte</i>
IRRITATION	énervement ; colère ; <i>agacement, mécontentement</i>	
MÉCONTENTEMENT	insatisfaction , <i>déplaisir, irritation</i>	<i>déception</i> ; <i>frustration</i>
MÉPRIS	dédain ; <i>irrespect</i> ; “lit” <i>mésestime</i>	<i>dégoût</i>
PEINE	mal ; <i>chagrin</i> ; <i>tristesse</i>	<i>déprime</i> ; <i>gêne, malaise</i>
PEUR	<i>crainte</i> ; “fam” <i>trac</i> ; “pop” <i>trouille</i>	
RANCUNE	<i>rancoeur, ressentiment</i>	
REGRET#1.1	<i>douleur</i> ; <i>tristesse</i>	<i>nostalgie</i>
REGRET#1.2	remords ; <i>repentir</i>	
REMORDS	culpabilité ; regret#1.2 , <i>repentir</i>	
REPENTIR	remords ; <i>regret#1.2</i>	
RESPECT#1	admiration ; <i>égard, estime, consideration</i>	

Tableau 4 : Quasi-synonymes et sens proches des lexies visées

Les quasi-synonymes imprimés en **Arial black** proviennent de notre corpus. Les autres quasi-synonymes (qui n'épuisent pas le répertoire de quasi-synonymes des lexies considérées) ont été ajoutés pour montrer au lecteur la richesse lexicale à laquelle on fait face (et qui pose des difficultés pour la description.)

Noter que certains **sens proches** ci-dessous ne dénotent pas des sentiments : INTOLÉRENCE ‘attitude’, RÉPROBATION ‘attitude’, DIFFÉREND ‘situation’ et RÉVOLTE ‘action’. (En effet, une différence importante entre quasi-synonymes et sens proches est que les premiers, mais pas les seconds, ont nécessairement la même étiquette sémantique ou, du moins, la même étiquette mère).

3.2 Les résultats par participant

Comme le montre le Tableau 5 ci-dessous, le meilleur résultat revient à un non francophone expert en lexicographie explicative et combinatoire, avec 17 réponses correctes sur 18. (La seule lexie que ce participant n'a pas reconnue est PEINE, la plus « française » des lexies de notre corpus.) Le pire résultat – une réponse correcte sur 18 – a été obtenu par un francophone sans connaissances en linguistique ; l'autre participant non initié à la linguistique a eu le second résultat le moins réussi, avec seulement 3 réponses correctes.

Il existe donc, comme nous nous attendions, une forte corrélation entre le taux de réussite en reconnaissance des lexies et le niveau de connaissances en linguistique que possèdent les participants. Voici un exemple simple de la façon dont les connaissances en linguistique peuvent guider les choix dans la tâche de reconnaissance. La maîtrise de la notion d'étiquette sémantique d'une lexie L (≈ genre prochain de L) permet de ne pas confondre des lexies de sens proche, qui, comme nous l'avons indiqué plus haut, ne peuvent pas porter la même étiquette sémantique. C'est le cas, par exemple, des lexies INDIGNATION, qui est une instance de sentiment, et RÉVOLTE, qui, elle, est une instance de acte.

⁵ L'acception de SCANDALE qu'on vise ici peut être illustrée par l'exemple suivant : *Il l'a épousée au grand scandale de sa famille.*

	Incorrectes			Correctes	Aucune	Francophone	Connaissances en linguistique
	QSyn	SProche	“Off-mark”				
[9]	0	0	0	17	1	NON	√/
[1]	1	0	5	12	0	√	√/
[3]	3	3	3	9	0	√	√
[7]	2	1	6	9	0	√	√
[2]	3	4	0	8	3	√	√
[8]	4	2	4	7	1	√	√/
[4]	5	5	2	5	1	√	√
[5]	4	4	6	3	1	√	NON
[6]	2	1	11	1	3	√	NON
	24	20	37	71	10		

Tableau 5 : Distribution des réponses par participant

4 Modification des définitions suite au test de reconnaissance

Nous commencerons par indiquer la définition corrigée de la lexie INDIGNATION, en expliquant comment les indications du test nous ont aidé à déceler les lacunes de la version originale de la définition, pour donner ensuite les versions finales des définitions de trois lexies quasi-synonymes de notre corpus : celles de REMORDS, REPENTIR et REGRET#1.2 (ces définitions n’ont été que légèrement modifiées par rapport à leur version initiale).

- INDIGNATION bis

~ DE LA personne X CONTRE acte/comportement Z DE LA personne Y ≡

‘Très fort sentiment négatif de _{personne}X envers _{acte/comportement}Z de _{personne}Y
causé par le fait suivant:

_{personne}X pense que _{acte/comportement}Z de _{personne}Y
viole intensément les principes de morale ou de justice de _{personne}X ;
ceci cause chez _{personne}X le désir de faire quelque chose
pour empêcher que des _{acte/comportement}Z semblables se reproduisent.

Ce sentiment est le sentiment que les gens ont normalement dans des situations semblables’.

Off-mark : CRIME ; JUSTICE || REGRET ; DÉGOÛT

SProche : RÉPROBATION ‘attitude’ ; RÉVOLTE ‘action’

Mis à part le fait qu’il s’agisse de sentiments déplaisants, il est difficile de trouver un lien entre les lexies REGRET et DÉGOÛT et la lexie visée (le fait qu’on puisse regretter les choses qui provoquent l’indignation ou en être dégoûté n’est pas lexicographiquement, ou même pragmatiquement, pertinent). Le lien entre CRIME et INDIGNATION est plus clair, le crime étant une des causes possibles d’indignation et une des instances possibles de l’actant Z de la lexie correspondante. Il en va de même pour le lien entre JUSTICE et INDIGNATION, car l’actant Z de cette dernière lexie est un acte ou un comportement contraire aux principes de la justice. (Ceci veut dire que, probablement, nous aurions dû traiter CRIME et JUSTICE comme des sens proches d’INDIGNATION.)

Le lien entre INDIGNATION et RÉPROBATION s’établit par la composante ‘Z(Y) est très mauvais moralement/socialément’, commune aux définitions des deux lexies : ce qui indigne, à cause de son inacceptabilité sociale et morale, et aussi répréhensible. Finalement, INDIGNATION est liée au RÉVOLTE par la composante ‘X veut faire qqch pour empêcher Z’; cette composante reflète, justement, le caractère dynamique de l’indignation, qui est un sentiment qui pousse à l’action.⁶

La première version de notre définition d’INDIGNATION ne tenait pas suffisamment compte de ces liens et était donc trop vague.

⁶ Cf. la définition d’INDIGNATION du *Petit Robert* : ‘sentiment de colère que soulève une action qui heurte la conscience morale, le sentiment de la justice’ → RÉVOLTE. (En passant, nous ne pensons pas qu’INDIGNATION soit une instance de COLÈRE, mais ne pouvons justifier notre point de vue ici, à cause des contraintes d’espace.)

- Série REMORDS ~ REPENTIR ~ REGRET#1.2

REMORDS (*Il avait dit au cours du procès n'éprouver aucun remords pour les attentats.*)

~ DE L'individu X À PROPOS DE L'action Y(X) ≡

'Fort sentiment négatif de individu X causé par le fait suivant:

individu X a/n'a pas fait action Y, ce qui est mauvais;

individu X est conscient du fait que individu X n'aurait pas dû/aurait dû faire action Y;

ceci cause chez individu X un mal moral.

Ce sentiment est le sentiment que les gens ont normalement dans des situations semblables.'

REPENTIR (*À son réveil, son premier sentiment fut un vif repentir de sa conduite.*)

~ DE L'individu X À PROPOS DE L'action Y(X) ≡

'Sentiment négatif de individu X causé par le fait suivant:

individu X a/n'a pas fait action Y, ce qui est mauvais;

individu X est conscient du fait que individu X n'aurait pas dû/aurait dû faire Y;

ceci cause chez individu X le désir de ne pas faire dans l'avenir une action Y semblable et de communiquer aux autres ce désir.

Ce sentiment est le sentiment que les gens ont normalement dans des situations semblables.

REGRET#1.2 (*Je n'ai aucun regret d'avoir fait le saut en politique*)

~ DE LA personne X À PROPOS DU fait Y ≡

'Sentiment négatif de personne X causé par le fait suivant:

X voudrait

que fait Y n'ait pas eu lieu ou que personne X ait réagi différemment à propos de fait Y

ceci est causé chez personne X par le fait que personne X voit les conséquences négatives de fait Y.

Ce sentiment est le sentiment que les gens ont normalement dans des situations semblables.'

Pour clore cette section, nous aimerions proposer un outil lexicographique qui permettrait de mieux distinguer des lexies sémantiquement apparentées, en l'occurrence, les lexies dénotant des sentiments. Il s'agirait de trouver, pour chaque sentiment, une expression française qui le « représente » le mieux : quelque chose que les gens disent effectivement lorsqu'ils ressentent le sentiment en cause. Ces clichés, qui ne correspondent pas forcément à des composantes de la définition de la lexie décrite, devraient figurer dans l'article de dictionnaire de cette dernière, où ils seraient décrits au moyen d'une fonction lexicale (non standard) de type :

{ce que dit X pour exprimer son [nom de sentiment]}⁷

Voici, à titre d'illustration, les valeurs de la fonction en cause pour les trois lexies dont les définitions viennent d'être données.

REMORDS : *Comment j'ai pu ?*

REPENTIR : *Je ne le ferais jamais plus !*

REGRET#1.2 : *Si seulement je n'avais pas fait cela !*

Il existe, en outre, un cliché commun aux trois lexies (ce qui reflète bien leur nature de quasi-synonymes) : *Je n'aurais pas dû faire cela.*

5 Conclusion

La plupart des participants au test présenté ci-dessus ont eu des difficultés avec la reconnaissance des lexies à partir des définitions proposées, ce qui indique que celles-ci n'étaient pas encore adéquates en ce qui concerne leur pouvoir distinctif.

Cette situation s'explique en partie par les faits suivants :

- Les lexies dont les définitions ont été testées appartiennent à un champ sémantique difficile (cf. note 2 ci-dessus). Notons aussi que certaines de ces lexies appartiennent au vocabulaire peu courant (par ex., RANCUNE et REPENTIR) et que d'autres ont été présentées avec leur « partenaire » au sein du même

⁷ Les fonctions lexicales sont des outils formels que la LEC utilise pour la description des relations lexicales, en particulier des collocation ; voir, par exemple, Wanner (ed.), 1996.

vocable (HAINE#1.a et HAINE#1.b ; REGRET#1.1 et REGRET#1.2 ; HOSTILITÉ#1 et HOSTILITÉ#2), ce qui a créé des difficultés additionnelles pour des participants non initiés à la linguistique (peu habitués aux faits de polysémie).

- Le test lui-même a été difficile, étant donné qu'on n'a offert aux participants qu'une liste des définitions à lire, sans indiquer les lexies « candidats » possibles.
- Les connaissances lexicales sont essentiellement floues, de sorte qu'on peut se demander si les gens connaissent vraiment le sens des lexies hors contexte. Elles sont aussi inconscientes, si bien qu'une recherche consciente de l'expression pour un sens donné nécessite un certain entraînement.

À cela il faut ajouter des traits individuels des participants, notamment leur niveau du français, leurs capacités générales et leur connaissances en linguistique. (Certains participants ne seraient pas capables de reconnaître les lexies correspondantes même si on leur présentait des définitions « idéales ».)

L'exercice présenté dans cet article a permis de mettre en évidence les deux points suivants, l'un important pour le lexicographe et l'autre pour l'enseignant de langue.

- L'utilité de considérer les sens proches de la lexie L pour laquelle on veut proposer une description lexicographique. Une comparaison entre L et ses sens proches met en évidence les composantes saillantes du sens de L et permet ainsi un premier « dégrossissage » efficace, après quoi on peut procéder à une comparaison avec les quasi-synonymes de L, pour effectuer le *fine tuning* de la définition.
- L'utilité des connaissances SUR la linguistique pour l'acquisition des connaissances DE la langue. Il est indicatif que le meilleur score du test revient à un non francophone ayant des connaissances expertes en linguistique.

Pour ce qui est du travail à venir, il serait intéressant, d'une part, de reprendre le test de reconnaissance avec les définitions corrigées et de vrais apprenants du français langue seconde et, d'autre part, de tester l'effet de la didactisation des définitions (qui n'a été que brièvement mentionnée dans cet article) sur la reconnaissance de lexies correspondantes.

Remerciements

J'aimerais remercier Igor Mel'čuk pour ses commentaires sur une version préliminaire de cet article. Je voudrais aussi reconnaître l'aide financière octroyée au projet *Dire Autrement* par le Conseil Canadien de recherche en sciences humaines (subvention de recherche n° 410-2005-0177).

Références

- Apresjan, V. & Apresjan, Ju. (1995). Metafora v semantičeskom predstavlenii èmocii. In: Apresjan (1995), *Integral'noe opisanie jazyka i sistemnaja leksikografija*. Moskva: Jazyki russkoj kul'tury, pp. 453-465.
- Apresjan, Yu., Djačenko, P., Lazurski, A. & Tsinman, L. (to appear). O komp'juternom učebnike leksiki russkogo jazyka. *Russkij jazyk v naučnom osveščeni*.
- Hamel, M.-J. & Miličević, J. (2007). Analyse d'erreurs lexicales d'apprenants du FLS : démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage. *Revue canadienne de linguistique appliquée*, 10/1, pp. 25-45.
- Iordanskaja, L. & Mel'čuk, I. (1999). Semantics of Two Emotion Verbs in Russian: BOJAT'SJA '[to] be afraid' and NADEJAT'SJA '[to] hope'. *Australian Journal of Linguistics*, 10:2, pp. 307-357.
- Mel'čuk, I. (2006). Explanatory-Combinatorial Dictionary. In: Sica, G. (ed.). *Open Problems in Linguistics and Lexicography*. Monza: Polimetrica Publisher; 225-355.
- Mel'čuk, I. et al. (1984-1988-1992-1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I-IV*. Montréal: Presses de l'Université de Montréal.
- Mel'čuk, I., Clas, A. & Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Leuven-la-neuve: Duculot.
- Miličević (1997). Étiquettes sémantiques dans un dictionnaire formalisé de type « Dictionnaire explicatif et combinatoire ». Mémoire de maîtrise non publiée. Montréal : Université de Montréal.

- Milićević, J. (2008). Structure de la définition lexicographique dans un dictionnaire d'apprentissage explicatif et combinatoire. In : Bernal, E. & DeCesaris, J., eds., *Proceedings of the XIII EURALEX International Congress*. Barcelona, 15-19 July, 2008. Barcelona: University Institute for Applied Linguistics, Pompeu Fabra University: 551-561.
- Milićević, J. & Hamel, M.-J. (2007). Un dictionnaire de reformulation pour les apprenants du français langue seconde. Chevalier, G, Gauvin, K. & Merkle, D., eds., In : *Les apports de la sociolinguistique et de la linguistique à l'enseignement des langues en contexte plurilingue et pluridialectal*. *Revue de l'Université de Moncton*, numéro hors série 2007; 145 -167.
- Polguère, A. (2003). Étiquetage sémantique des lexies dans la base de données DiCo. *Traitement Automatique des Langues* (T.A.L.), 44/2: 39-68.
- Popovic, S. (2003). *Paraphrasage des liens de fonctions lexicales* [mémoire de maîtrise]. Département de linguistique et de traduction, Université de Montréal.
- Wanner, L., ed. (1996). *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam/Philadelphia: Benjamins.
- Wierzbicka, A. (1992). *Semantic, Culture and Cognition. Universal Human Concepts in Culture-Specific Configurations*. Oxford University Press.
- Wierzbicka, A. (1999). *Emotions Across Languages and Cultures*. Cambridge: Cambridge University Press.