

Transparences lexicales entre langues voisines

ÉRIC CASTAGNE
CIRLLLEP (EA3794)

Université de Reims Champagne-Ardenne (France)

RÉSUMÉ : Dans le cadre du programme InterCompréhension Européenne dont l'objectif est d'analyser de manière empirique, de développer et d'optimiser l'approche et la gestion de l'information et de la connaissance en multiples langues (pour l'instant, français, espagnol, italien, portugais, roumain, anglais, allemand et néerlandais), nous avons considéré qu'avant tout, nous devons exploiter les connaissances déjà à disposition au moins dans notre langue maternelle pour lire dans une autre langue. La première des connaissances repose sur la reconnaissance de transparences entre les langues. Mais que sont les transparences entre plusieurs langues ? L'objet de cette communication sera de définir la notion de "transparence", et plus particulièrement la notion de "transparence lexicale" entre plusieurs langues voisines – espagnol, italien, portugais – ou apparentées – anglais, allemand, néerlandais – pour un lecteur francophone, de proposer une typologie des degrés de la "transparence lexicale" – en terme de mots "directement transparents", mots "opaques", mots "indirectement transparents" –, et de montrer comment faciliter l'accès aux "mots indirectement transparents" afin que leur reconnaissance, accompagnant celle des "mots directement transparents", rende suffisamment conséquentes les indications sémantiques pour permettre la compréhension des textes et des discours.

ABSTRACT: Within the framework of the European InterComprehension programme the objective of which is to empirically analyse, develop and optimise the approach and management of information and knowledge in several languages (French, Spanish, Italian, Portuguese, Romanian, English, German, Dutch), we considered that we must in the first place take advantage the knowledge of our mother tongue to read in another language. The first element of this knowledge relies on the recognition of transparencies between the languages. But what are the transparencies between several languages? The object of this paper is threefold: we will first define the concept of "transparency", and more particularly the concept of "lexical transparency" between several close languages - Spanish, Italian, Portuguese - or related - English, German, Dutch - for French-speaking readers. Then we will propose a typology of the degrees of "lexical transparency" - in term of "directly transparent" words, "opaque" words, "indirectly transparent" words -, and finally we will show how to facilitate access to the "indirectly transparent words" so that their recognition, together with that of the "directly transparent words", gives enough semantic indications so as to allow the comprehension of both written and oral texts.

1. LA NOTION D'INTERCOMPRÉHENSION

Nous avons dit que l'intercompréhension fonctionnelle entre langues était le fait de comprendre une ou plusieurs langues sans la ou les parler. La performance d'une telle communication plurilingue est optimale dans les situations où les personnes maîtrisent déjà les deux langues utilisées dans la communication (par exemple, un francophone connaissant l'anglais et un anglophone connaissant le français). Cette pratique est également très fonctionnelle dans les situations où les personnes ne maîtrisent que les compétences de compréhension(s) écrite et/ou orale des autres langues.

L'objectif du programme InterCompréhension Européenne (ICE) est de mener dans un esprit collaboratif et pluridisciplinaire une réflexion et des recherches ouvertes, concrètes et efficaces sur l'intercompréhension des peuples entre eux, et de répondre avec efficacité aux besoins concrets actuels, en développant et en optimisant l'approche et la gestion de l'information et de la connaissance en multiples langues.

Elaboré après des recherches fondamentales et appliquées, le programme ICE propose de développer une intercompréhension très pratique qui facilite la "communicance", d'une part dans le domaine de l'acquisition de l'information parce qu'elle permet d'une façon très pratique la veille informationnelle et technologique multilingue, et d'autre part dans le domaine de la communication de l'information parce qu'elle apprend à communiquer synthétiquement, clairement et justement, ce qui est capital dans les échanges industriels et économiques, sans parler du fait que cette facilité abaisserait les coûts à différents niveaux.

Pour atteindre cet objectif, l'équipe du programme ICE a développé une méthodologie et est prête à élaborer un DVD intégrant des formations qui visent, non pas à façonner des spécialistes en langues ou en traduction, mais à offrir des compétences de compréhension écrite et orale en plusieurs langues européennes (pour l'instant, espagnol, français, italien, portugais, roumain, anglais, allemand, néerlandais), et de nouvelles compétences d'expression écrite et orale dans sa langue maternelle pour une meilleure communication plurilingue.

Un schéma de large dissémination est projeté : un peu partout, des chercheurs œuvreront avec divers groupes, dans plusieurs régions d'Europe et du monde (où les mêmes langues sont parlées). Pour finir, tout se reliera en une grande toile d'intercompréhension : langues romanes, langues germaniques, langues slaves, langues scandinaves, ...

Cette communication portera sur une des dimensions en jeu dans les processus de compréhension : la dimension lexicale.

Elle s'organisera en deux étapes, l'une tentant de définir la notion de transparence lexicale et de ses variantes, l'autre vérifiant cette fonctionnalité par la quantification des phénomènes dans le cadre de l'intercompréhension entre francophones et anglophones.

2. DÉFINITION DE LA NOTION DE TRANSPARENCE LEXICALE ET DE SES VARIANTES

2.1. Mots "transparentes"

Nous appelons "séquences transparentes" les séquences, qu'elles soient graphiques, phoniques, morphologiques, ou syntaxiques, qui peuvent être jugées lisibles par un lecteur "débutant" dans une langue étrangère (LE) à partir des seules connaissances extraites dans sa langue maternelle (LM).

- a)
- | | | |
|--------------------------------|-------------------|------------------------|
| | Universidade (PT) | |
| | Universidad (ES) | |
| Un gruppo di ricercatori dell' | Università (IT) | tedesca di Bamberg ... |
| | Université (FR) | |
| | University (GB) | |
| | Universiteit (NL) | |
| | Universität (DE) | |
- b)
- | | | |
|--------|----------------|--|
| | turistas (PT) | |
| | turistas (ES) | |
| | turiste (IT) | |
| | touristes (FR) | |
| | tourists (GB) | |
| Hordes | toeristen (NL) | zullen in 2009 flaneren door het centrum ... |
| | Touristen (DE) | |

Dans cette communication, nous nous intéresserons plus particulièrement aux mots lexicaux dans une perspective analytique. Et nous appellerons "mot transparent" tout terme lexical, présent dans un texte en une LE donnée qui d'une part possèdera une forme qui suggère un terme lexical en LM de forme (écrite ou orale) quasi identique, et qui d'autre part se trouve posséder une signification quasi-identique dans les deux langues.

On parle alors de quasi identité formelle et de quasi identité sémantique. La démarche est donc d'esprit typiquement sémasiologique et le mot transparent fonctionne comme icône.

2.2. Mots "non transparents"

Nous appellerons "mot non transparent" tout terme lexical présent dans un texte en LE qui peut être jugé illisible par un lecteur "débutant" dans cette même LE à partir des connaissances extraites dans sa LM.

Parmi les mots "non transparents", les observations faites dans les séances empiriques ICE ont permis de dégager deux classes distinctes de "mots non transparents" :

2.2.1. Mots "opaques"

Les "mots opaques", qui sont des termes lexicaux (présents dans un texte en une LE donnée) qui possèdent une forme qui, soit ne suggère aucun terme lexical en LM, soit suggère un terme lexical en LM mais qui possède une signification autre.

- Ceux dont la forme ne suggère aucun terme lexical en LM :

c)	
IT	FR
Risalire (a)	remonter (à)
mela	pomme
ES	FR
Empezar	commencer
Melocoton	pêche
NL	FR
Zolder	grenier
Kunst	art
DE	FR
(Aus)wählen	choisir
Bedeuten	signifier

- Et ceux dont la forme suggère un terme lexical en LM mais qui possède une signification autre ; dans ce cas, il est convenu traditionnellement d'appeler ce genre de termes faux-amis.

d)		
IT	faux-ami	FR
alcune	aucune	quelques
ES	faux-ami	FR
salir	salir	sortir
NL	faux-ami	FR
de	de	le, la

Il faut préciser immédiatement que cette question des mots opaques, et en particulier des faux-amis, n'est pas un problème fondamental dans l'objectif qui nous intéresse d'une part parce que nous travaillons sur la compréhension et non sur l'expression (donc cela présuppose que la version originale est exacte), et d'autre part parce que, si l'on bloque sur un mot opaque ou si l'on part sur une fausse piste, le contexte nous remet très rapidement sur le droit chemin.

e)

IT

i risultati di alcune indagini commissionate dal governo indicano che molti bambini non possiedono le nozioni più elementari di matematica o di lingue inglese e ...

FR

les résultats de quelques enquêtes commandées par le gouvernement indiquent que ...

f)

DE

Ein genialer Forscher ohne Dokortitel revolutionierte im Jahr 1905 in einsamer Heimarbeit das Weltbild der Physik.

FR

Un génial chercheur ?AVEC? titre de Docteur révolutionna en 1905 dans un travail personnel à la maison l'image de la physique ???

Un génial chercheur SANS titre de Docteur révolutionna en 1905 dans un travail personnel à la maison l'image de la physique.

Il y a donc deux sous-types de mots opaques : selon le cas, on parle d'altérité formelle, ou de quasi-identité formelle avec altérité sémantique.

2.2.2. Mots "semi-transparents"

Les "mots semi-transparents" sont les mots qui ne sont pas jugés immédiatement lisibles par un lecteur "débutant" dans une LE à partir des connaissances extraites dans sa LM, mais ils comportent un (des) indice(s) permettant d'en comprendre la signification au moyen d'un "mot-relais". Ici aussi semblent se dégager deux sous-types généraux :

- 1er sous-type :

g)

mot en LE

cena (ES/IT)

parar (PT/ES)

encender (ES)

feliz (PT/ES)

umbrella (GB)

book (GB) / boek (NL)

Wagen (DE)

trinken (DE)

mot-relais

cène [rel.] (FR)

parer [un coup] (FR)

incendier (FR)

félicité (FR)

ombrelle (FR)

bouquin (FR)

wagon (FR)

trinquer (FR)

mot en LM

repas (FR)

arrêter (FR)

allumer (FR)

heureux (FR)

parapluie (FR)

livre (FR)

voiture (FR)

boire (FR)

Le terme de la LE suggère formellement un terme de la LM dont la signification, même si elle est autre, est tout de même accessible en tenant compte des autres significations présentes dans le contexte : le terme fonctionne alors comme une sorte d'approximation appelant un dépassement par substitution avec un terme ajusté au contexte. Le "mot-relais" est dans ce cas un mot réel de la LM.

- 2ème sous-type (Tous les exemples des langues romanes sont extraits de la méthode Eurom4, ceux des langues germaniques du programme ICE) :

h)							
PT	ES	IT		mot imaginaire	FR		
talhar	tallar	tagliare		*tallier	tailler		
vinha	viña	(vigna)		*vigna	vigne		
GB	NL	DE		mot imaginaire	FR		
(billion)	miljard [lj]	(Milliarde)		*miliard	milliard [lj]		
(million)	miljoen	(Million)		*milion	million		
GB	NL	DE		mot imaginaire	FR		
in	aan	an		*an	en		
board	boord	(bord)		*bor(d)	bord		
a(n)	een	ein		*ein	un		
PT	ES	IT		mot imaginaire	FR		
bomba[b]	bomba[b]	(pompa[p])		*pompa	pompe[p]		
GB	NL	DE		mot imaginaire	FR		
flam	vlam	Flamme		*flam	flamme		
fruit	vrucht	Frucht		*fruc(h)t	fruit		
PT	ES	IT		mot imaginaire	FR		
carne[k]	carne [k]	carne [k]		*charne	chair		
					(charnel)		
GB	NL	DE		mot imaginaire	FR		
(braise[z])	braten [t]	Braden [d]		*braser	braiser [z]		
break[k]	breken [k]	(brechen [ʃ])		*brécher	ébrécher [ʃ]		
					casser		
PT	ES	IT	GB	NL	DE	mot imaginaire	FR
haver	haber[b]	avere	have	hebben	haben	*haver	avoir

Le terme de la LE ne suggère, de par sa forme stricte, aucun terme de la LM, mais peut en suggérer un à condition de faire varier sa forme (ce qui nous éloigne du mot originel) en affaiblissant ici ou là la précision de la réalisation de tel ou tel de ses graphèmes, et/ou de tel ou tel de ses phonèmes jusqu'à faire apparaître un *terme imaginaire* de la LE qui fonctionne alors comme un mot transparent. Le mot-relais est dans ce cas un terme imaginaire de la LE.

2.2.3. Conclusion partielle

Dans la situation du premier sous-type, le mot-relais est un mot réel de la LM qui ne convient pas du point de vue sémantique. Mais intervient alors un réajustement qui conduit à un autre mot réel qui n'est pas, lui, en relation de quasi-identité formelle avec le mot original de la LE. Ce réajustement de signification, opéré *dans* la LM, doit être considéré comme une sorte de médiation sémantique (dans la LM).

Dans la situation du second sous-type, le mot-relais est un mot imaginaire de la LE dont la forme déclenche une transparence. Ici il s'agit d'une manipulation de forme, opérée *dans* la LE, qui doit être considérée comme une sorte de médiation formelle (dans la LE).

Ainsi est capitale l'observation que les médiations sémantiques sont effectuées dans la LM et font partie de la compétence native alors que les médiations formelles dans la LE n'exigent aucune compréhension de la LE. Dans les deux cas, on parlera de transparence indirecte, en distinguant donc deux types de transparence indirecte.

2.3. Proposition de typologie

La classification des termes lexicaux d'une LE qui vient d'être effectuée peut se schématiser dans le tableau suivant. On ne prétendra pas que la typologie est stricte. Cela pourra être une question d'appréciation (personnelle en fonction de ses habiletés) de considérer que certains mots sont à la frontière de 2 types. Suivant la LE, des mots de la même série pourront être considérés comme transparents ou semi-transparentes. D'où l'importance de travailler simultanément sur plusieurs langues.

mots transparents	mots non-transparentes			
	mots semi-transparentes		mots opaques	
<i>transparence directe</i>	<i>transparence indirecte</i>		<i>opacité</i>	
	Sous-type 1	Sous-type 2	Sous-type 1	Sous-type 2
	le relais effectue une médiation sémantique au sein de la LM (faux-amis solubles)	le relais effectue une médiation formelle au sein de la LE	ne suggère aucun terme lexical en LM	suggère un terme lexical en LM mais qui possède une signification autre (faux-amis non solubles hors contexte)
universidade (PT) universidad (ES) università (IT) université (FR) university (GB) universiteit (NL) Universität (DE)	umbrella(GB) ombrelle (F) parapluie (F)	miljoen (NL) *milion million (F)	zolder (NL) grenier (F)	alcune (I) ?aucune quelques (F)

3. QUANTIFICATION DES PHÉNOMÈNES : L'EXEMPLE DU COUPLE FRANCO-ANGLAIS

3.1. Mots transparents vs mots non transparents

francophone devant l'anglais britannique LE ⁹⁶				anglophone devant le français de France LE ⁹⁷			
%	transparent	non transparent	total	%	transparent	non transparent	total
V ⁹⁸	29	71	100	V ⁹⁹	47	53	100
N ¹⁰⁰	48	52	100	N ¹⁰¹	43	57	100

Si l'on observe les résultats généraux de transparence de l'anglais pour un francophone, on observe une majorité de non transparence, même si dans le cas des Noms il y a un quasi équilibre.

Si l'on observe les résultats généraux de transparence du français pour un anglophone, on observe un certain équilibre dans la transparence, toutes catégories grammaticales confondues, malgré une légère prédominance de la non transparence.

Si l'on observe plus dans le détail, il y a plus de Verbes français transparents pour un anglophone que de Verbes anglais transparents pour un francophone (47 % / 29 %).

En revanche, il y a de manière moins nette plus de Noms anglais transparents pour un francophone que de Noms français transparents pour un anglophone (48 % / 43 %).

3.2. Intégration de la notion de semi-transparence dans le comptage

Suite à une consultation en séance d'observation, les résultats intégrant la distinction mots opaques et mots semi-transparentes sont :

francophone devant l'anglais britannique LE					anglophone devant le français de France LE				
%	transparent	semi-trans.	opaque	total	%	transparent	semi-trans.	opaque	total
V	29	23	48	100	V	47	20	33	100
N	48	17	35	100	N	43	23	34	100

La transparence indirecte représente approximativement 30 % des mots non transparents anglais pour les francophones, et 40 % des mots non transparents français pour les anglophones, aussi bien pour les Verbes que pour les Noms.

96 Cf. M. Caure, 2005

97 Cf. A. Chantegrail, 2005

98 Les 700 verbes les plus fréquents de l'anglais britannique

99 Plus de 420 verbes parmi les plus fréquents

100 Les 1100 noms les plus fréquents de l'anglais britannique

101 Plus de 620 noms parmi les plus fréquents

La transparence indirecte représente approximativement 20 % des mots dans les deux situations, c'est-à-dire environ un mot sur cinq.

3.3. Transparences directe et indirecte vs opacité réelle

francophone devant l'anglais britannique LE				anglophone devant le français de France LE			
%	tr. directe et indirecte	opacité	total	%	tr. directe et indirecte	opacité	total
V	52	48	100	V	67	33	100
N	65	35	100	N	66	34	100

La transparence potentielle est toujours supérieure à la moitié des mots dans les 2 situations. Elle atteint même régulièrement les 2/3 dans la situation de l'anglophone face au français.

3.4. Bilan

La prise en compte des phénomènes de transparence indirecte fait aussi passer la reconnaissance par un francophone des Verbes anglais de 1/3 à plus de la moitié, et la reconnaissance des Noms anglais de la moitié à plus des 2/3.

La prise en compte des phénomènes de transparence indirecte fait aussi passer la reconnaissance par un anglophone des Verbes et des Noms français de moins de la moitié au 2/3.

La prise en compte non seulement de la transparence indirecte mais aussi du poids des Noms et des Verbes dans la structure sémantique des énoncés fait apparaître que dans le sens français > anglais comme dans celui anglais > français, l'information (directement ou indirectement accessible) représente souvent les 2/3 des informations nécessaires à la compréhension.

On connaît bien la proximité lexicale historiquement entretenue entre ces deux langues, et on ne prétend pas que les mêmes valeurs pourraient être trouvées en considérant d'autres langues germaniques que l'anglais.

4. CONCLUSION

Cette tentative de typologie de la transparence lexicale entre plusieurs langues et son expérimentation sur le couple franco-anglais présentent déjà des résultats encourageants pour envisager une intercompréhension fonctionnelle entre l'anglais et le français. Mais il faut déjà envisager une amélioration du seuil d'intercompréhension.

D'une part, dans le cadre de la transparence indirecte avec relais sémantique dans la langue maternelle, on sait que les effets de contexte ne sont pas négligeables, sans parler de l'importance de la connaissance du

lexique de la langue maternelle, y compris dans le traitement de la référence. Dans le cadre de la transparence indirecte avec relais formel dans la langue étrangère, il faut prendre en compte les effets de "flou ordonné" morphologique et l'importance de la connaissance des possibilités d'exploitation de proximités articulatoires (avec une systématisation très rapide).

D'autre part, favorisent la compréhension des facteurs extra-lexicaux comme la mémorisation-capitalisation de chemins inférentiels (voir Tyvaert dans ce même ouvrage), le travail concomitant sur les structures syntaxiques (inférences syntactico-sémantiques), le remplacement de l'analyse des facilités d'ordre lexical dans la démarche générale (lire non pas les mots mais les séquences avec une tolérance autant du point de vue de l'exactitude formelle que du point de vue de la correction sémantique), les renforcements induits par le travail simultané sur trois langues (étrangères) apparentées et l'exploitation de toutes les ressources linguistiques à disposition (langue maternelle, une ou plusieurs langues étrangères).

La pratique de la forme de plurilinguisme que nous préconisons, que l'on qualifie souvent d'asymétrique, (chacun parle sa langue maternelle et comprend plusieurs langues apparentées ou voisines), nourrit et dynamise continuellement le système de reconnaissance des liens de parenté et de voisinage. Plus on pratique et plus on découvre des transparences là où dans un premier temps on ne les avait pas repérées.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Castagne Eric & Chartier Jean-Paul, (2007). "Modélisation de la formation d'éducateurs à l'intercompréhension de plusieurs langues : réflexions et pistes", in *Le Français Dans Le Monde*, n° spécial R/A, pp. 66-75
- Castagne Eric (sous la coordination de), (à paraître en 2007). *Les Enjeux de l'intercompréhension*, Coll. ICE, 2, Reims, éditions et presses universitaires de Reims, 347 p.
- Castagne Eric, (2004). *Intercompréhension européenne et plurilinguisme : propositions pour quelques aménagements linguistiques favorisant la communication plurilingue*, Aachen, ed. Shaker-Verlag, série éditions EuroCom.
- Caure Mélisandre, (2006). *Détermination et exploitation des transparences lexicales entre l'anglais et le français*, Mémoire de Master2 (sous la direction d'E. Castagne & J.-E. Tyvaert), Université de Reims Champagne-Ardenne, 99 p.
- Caure Mélisandre, (2005). *Correspondances lexicales entre deux langues voisines*, Mémoire de Master1 (sous la direction d'E. Castagne & J.-E. Tyvaert), Université de Reims Champagne-Ardenne.

- Chantegrail Aude, (2006). *Etude microsémantique des ressemblances lexicales entre le français et l'anglais*, Mémoire de Master2 (sous la direction de J.-E. Tyvaert & E. Castagne), Université de Reims Champagne-Ardenne, 174 p.
- Chantegrail Aude, (2005). *Lexical likeness between two languages*, Mémoire de Master1 (sous la direction de J.-E. Tyvaert & E. Castagne), Université de Reims Champagne-Ardenne.
- Laks Bernard, (à paraître en 2007). "De l'approximation dans la relation phonétique / phonologie", in Castagne (sous la coordination de) (à paraître en 2007).
- UNESCO, (2005). *Vers les sociétés de savoir* (rapport mondial de l'UNESCO), Paris, éditions UNESCO.