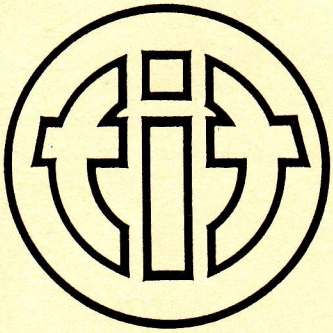
ISSN 1027-8087

TRANSLATIO

NOUVELLES DE LA



NEWSLETTER

RÉVUE TRIMESTRIELLE publiée par la
FEDERATION INTERNATIONALE DES TRADUCTEURS
International Federation of Translators
ONG en relations formelles de consultation de l'Unesco

Nouvelle série XVII (1998) No 3

Rédacteur en chef et éditeur responsable: Dr.R.HAESERYN
Heiveldstraat 245, B 9040 Sint-Amandsberg (Belgique)

Afgiftekantoor 9040 Sint-Amandsberg 1

TRANSLATIO
NOUVELLES DE LA FIT – FIT NEWSLETTER
Nouvelle série XVII (1998) – n° 3

SOMMAIRE – CONTENTS

ARTICLES

VIE DES ASSOCIATIONS MEMBRES ET AUTRES

VIE DES COMITES FIT et du CRE

CORE-ORIENTED LEXICAL TRANSLATION EVALUATION (COLTE)

JAHANGIRI, A.M.
Shiraz Azad University

0. Introduction

It is evident to anyone that words play an important role in every kind of faithful communication especially translation. The study of words can be linguistic, psycholinguistic, sociolinguistic, etc. In other words, it draws on a wide number of perspectives from which it can be viewed.

The attitude that translators have had towards the scientific study of words in translation has varied enormously during the past decades. Mastering the proper equivalent words of the two languages has long been considered essential for achieving fluency in translation. Translators often claim that their primary problem in translation is a lack of equivalent words in the target language. One way to classify words is the core/noncore distinction. In effect, all words are not of equal importance to a language user. Core vocabulary refers to those elements which are often the most basic, natural, and frequent words used by a native speaker of the given language (See "markedness" in Richards, J., Platt, J. & Weber, H., 1989). They are of much more frequency and use and are sometimes referred to as "nuclear vocabulary" in the related literature. One comes to the true conclusion that core words are of much more use than their noncore counterparts (See Ellis, 1985 for more details). As frequency counts show, round about 2000 vocabulary items constitute 90% of the totality of words used by native speakers of English (Celce-Murcia, M., & McIntosh, L., 1979).

I. The Application of Carter's Six Features : Some Extension

In a lucid explication of the point, Carter (1988) designed six tests to identify core vocabulary. He believes that core vocabulary items are char-

acterized by the following features (henceforth referred to as F1, F2, F3, ..., F6).

(1) One can easily find "clear antonyms" for them. For instance, in the lexical set "*dirty, filthy, grimy,* and *grubby*", the word *dirty* constitutes the core item because one can easily arrive at an antonym for it, namely *clean*, while this cannot be applied to *filthy, grimmy* and *grubby*.

The Core item . . → dirty ≠ clean
filthy
grimy
grubby

(2) Core words have a more extensive collocational span. That is, core words occur with more words than their noncore counterparts. The word *dirty* may be followed by *dog, man, trick, business, face, hands, streets, water*, etc. but the other members of the lexical set, the words *filthy, grimy* and *grubby* are not of the same collocational span as that of *dirty*. In simple words, they cannot be accompanied by the same number of words as that of the word *dirty*. For instance, one cannot derive a *filthy, grimy* or *grubby* man. They simply do not collocate. In this respect only the word *dirty* can precede *man*. Similarly, one cannot use the phrase *wash his/her filthy hands* (sic). He/She can only use the phrase *was his/her dirty hands*. The writer had some informants (native speakers) tick

Dirty		Filthy		Grimy		Grubby		
✓	dog		dog		dog		dog	
✓	man		man		man		man	
✓	trick		trick		trick		trick	
✓	business		business		business		business	
✓	face		face	✓	face	✓	face	
✓	hands		hands	✓	hands	✓	hands	
✓	streets	✓	streets		streets		streets	
✓	water	✓	water		water		water	
8/8		2/8		3/8		2/8		

Figure 1 : *The Relative Collocational Span Count*

the following as research subjects. They were required to tick the most natural sequence. They unanimously arrived at the following result. (This researcher-made test can be useful in the mathematical or quantitative measurement of the collocational span of several elements of a lexical set).

Care is to be taken not to claim more for the test than it deserves. The point must not be discussed in absolute terms. In this particular case "relativity" has a crucial meaning. In other words, the results mentioned above are drawn only with respect to the semantic realm which is composed of the elements *dog, man, trick, business, face, hands, street,* and *water*. That is, the word *dirty* has a more collocational span (i.e. it is used together with a greater number of the elements or words of the above-mentioned lexical realm) than the words *filthy, grimy,* and *grubby* whose collocational span count is equal to 2/8, 3/8 and 2/8, respectively. Nevertheless, the list does not seem to be exhaustive in any way. The above-mentioned notion of relativity can be best explained with respect to the same fact. Of course, the more elements the list has, the more accurately it measures the collocational span. For instance, a more accurate conclusion can be drawn if we include *family, windows, house, language* and *clothes*. Summing up we derive :

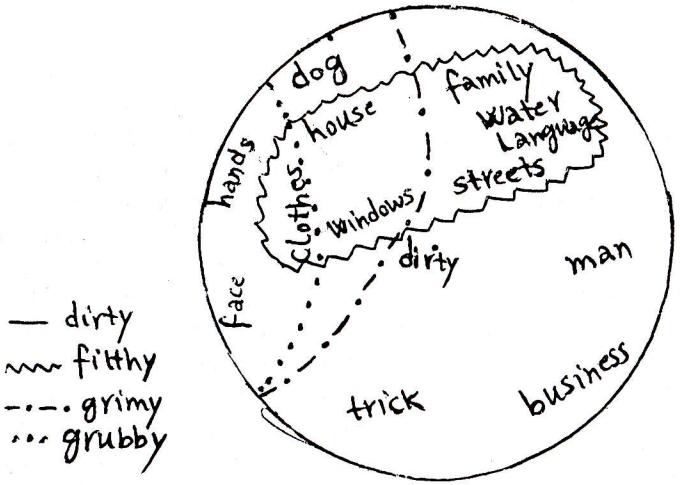
DIRTY	FILTHY	GRIMY	GRUBBY
13/13	7/13	6/13	3/13

The Relative Collocational Span Count

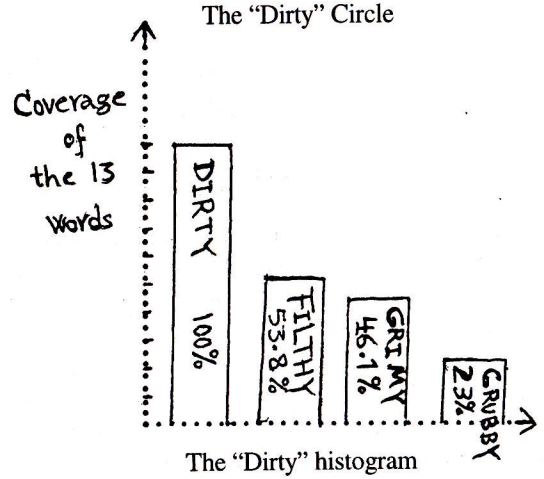
Relatively speaking, one may come to the conclusion that *dirty* has a more extensive collocational span (hence "the core item"). On the other hand, *grubby* can be said to be a "noncore item", because it is of the least degree of coreness.

Perhaps the coreness of *dirty* can be best represented by locating it at the centre. As the figure shows, the word *dirty* is more central to the English language than the words *filthy, grimy* and *grubby* (hence Core-Oriented Lexical Translation Evaluation or COLTE). The figure also shows that the commonality between the *filthy* set and the *grimy* set comprises of the

three words *clothes, windows,* and *house*. Using a mathematical notation we have [*Filthy* ∩ *Grimy* = 3]. The same thing can be applied to the other sets, as a result of which we have [Grubby ∩ Grimy = hands, face, and clothes = 3]



The "Dirty" Circle



The "Dirty" histogram

Figure 2

(3.) Core words are used to define the meanings of the other members of the lexical set. Carter states

> For example, in the set *snigger, grin, smirk, beam, smile,* all the words except *smile* can be defined by *smile* (the core item) plus an adverb. For example, *beam-smile happily; smirk = smile knowingly* (Carter 1988 : 10).

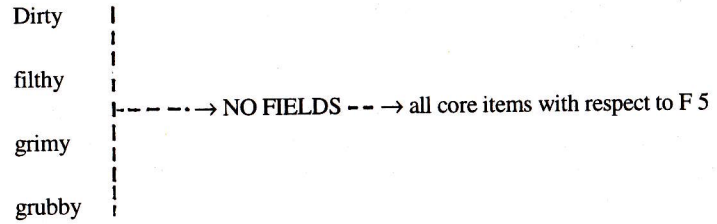
As applied to our case, if the careful reader looks the words *filthy, grimy* and *grubby* up in a good dictionary, he/she will find that they have all been defined with reference to the core word *dirty*.

(4.) "Core vocabulary items are those which do not carry especially marked connotations or associations" (Carter 1988 : 10).

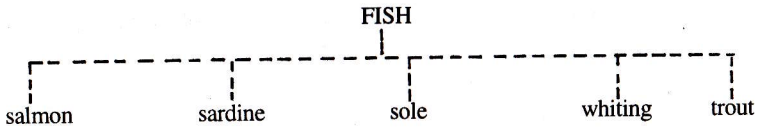
> dirty hands, *dirty mind*
> filthy boots, *filthy mind* (used figuratively)
> grimy hands, grimy windows, *grimy rascal*
> grubby hands, *grubby scandal*

As can be inferred, *dirty* cannot be considered as being the "core item". Therefore, features 1, 2 and 3 identify *dirty* as being the "core item" while feature 4 does not. Let's consider Carter's case. Carter (1988 : 10) states that " … in the set of words surrounding a core item such as *thin*, lexical items such as *skinny* or *slim* carry marked negative and positive associations respectively".

(5.) Scarcely any field can be attributed to core vocabulary items. That is, they do not belong to a particular field of discourse (Carter 1988). This test does not lend itself well to the case we had with the lexical set *dirty, filthy, grimy*, and *grubby*. Therefore, it cannot be used as a measure of coreness in this particular case. Of course, this is not to go so far as to say that *dirty* is not the core item, because it is, and this can be accounted for by the fact that it does not belong to a specific field of discourse. The other members of the set do not belong to a specific field of discourse either. Therefore, this makes the comparison invalid.



(6) "Core words are often superordinates" (Carter 1988 : 10). Therefore, in the set *salmon, sardine, sole, whiting* and *trout*, the superordinate item *fish* can stand for all the other members of the set.



Once again this feature of coreness does not apply to our case (here the lexical set *dirty, filthy, grimy* and *grubby*), because there is no clear-cut superordinate relationship.

By way of conclusion, one can infer the fact that using Carter's 6-feature model for determining core items the word *dirty* can be considered to be the core item with respect to features 1,2,3 and 5; that is 4 categories out of the 6 items in Carter's model (4/6).

2. What does all this have to do with lexical translation evaluation practice?

Lexical translation evaluation is done quantitatively using a scalar measure here. Corresponding values are obtained for the seemingly Persian and English equivalents. The numbers are derived with respect to the 6-item model developed by Carter (1988). A specific number is obtained to represent the degree of coreness of a given Persian/English equivalent. Then, the two numbers are compared. The less the difference between the

calculated numbers in Persian and English, the more they are likely to be used as equivalents.

3. Some Contrastive Lexical Analysis

Suppose one wants to find the best Persian equivalent among the lexical set *kasīf, nāpāk, ālūdeh,* and *cherkin*, for *dirty* (in the lexical set *dirty, filthy, grimy* and *grubby*). The comparison and application of Carter's six features in both Persian and English among the members of similar lexical fields leads to :

- The application of Feature 1

The first feature in Carter's model states that core words are associated by clear antonyms. Considering the English lexical set *dirty, filthy, grimy,* and *grubby*, the item "dirty" will soon remind one of the antonym "clear", while a clear antonym is too hard to be found for the other members of the lexical set *filthy, grimy* and *grubby.* The same analogy can be extended to the Persian lexical set *kasīf, nā-pāk* (not clean), *ālūdeh*, and *cherk-in* (dirt-like). *Kāsif* most often reminds a Persian speaker of *tamiz* (± clean), i.e. a clear antonym. This being the case, we derive 1 degree of coreness for *dirty* and *kasīf*, and a 0 degree of coreness (in fact it is equal to noncoreness) for the other members of the two lexical sets.

- The application of Feature 2

Feature 2 deals with collocational span; that is, we want to see whether the given word can be accompanied or followed by some other words or not. For instance, we can use the phrase "a dirty dog" but not "a filthy dog". For this section we used Persian and English informants to validate the results. Thirteen words were chosen to see if the members of both English and Persian lexical sets can follow / precede them. These were *dog, man, trick, business, face, hands, street, water, family, windows, house, language,* and *clothes*. If the two words cannot be put together in the intended language no value is obtained. Applying the notion to the English set we derive :

Dirty	*Filthy*	*Grimy*	*Grubby*
13/13	7/13	6/13	3/13
I	.53	.46	.23

In the case of the corresponding Persian lexical set we had informants tick the following table to find the collocational span with regard to the same 13 words. The following roughly-sketched generalization turned out to be of the highest frequency.

cherkin	ālūdeh	nāpāk	kasīf	The 13 words in Persian	
	✓		✓	"sag"	(dog)
		✓	✓	"ādam"	(man)
			✓	"hoqeh"	(trick)
			✓	"kār"	(business)
			✓	"sourat"	(face)
✓	✓		✓	"dasthaye"	(hands)
			✓	"xiābān"	(street)
	✓		✓	"āb"	(water)
		✓	✓	"xānevādeh"	(family)
			✓	"panjereh"	(window)
			✓	"xāneh"	(house)
		✓	✓	"zabān"	(language)
✓		✓	✓	"lebāshaye"	(clothes)

Figure 3 : Contrary to English, in Persian, adjectives follow the nouns they modify

That is,

kasīf	>	napāk	>	āludeh	>	cherkin
13/13		4/13		3/13		2/13
1		.3		.23		.15

where > means "having more coreness with respect to feature 2". The results are inserted in Figure 4 for later clarification purposes.

- The application of Feature 3

To see if the items *filthy, grimy,* and *grubby* are defined by the core word *dirty* we looked them up in the *Oxford Advanced Learner's Dictionary* (1989).

The three were all defined by the core word *dirty*. Adopting a scalar or quantitative approach for the determination of the degrees of coreness we derive 1 for *dirty* as the core item and 0 for *filthy, grimy*, and *grubby*. The same thing can be applied to the corresponding Persian lexical set *kasīf, nā-pāk, ālūdeh*, and *cherkin*. This was done with reference to Noi'n's Persian-Persian dictionary (Noi'n, 1983). We found that the members *nā-pāk, ālūdeh*, and *cherkin* have all been defined by *kasīf*. Following the first feature advocated by Carter for the determination of core items, one can refer to *kasīf* as the core item. As a result the relative degree of coreness would be set at 1 for *kasīf* and 0 for *nā-pāk, ālūdeh*, and *cherkin*. (For more information and a better representation of the procedure see Figure 4).

- The application of Feature 4

The application of Feature 4 shows that a word must not, necessarily, be labelled "core" considering all the 6 items in Carter's model. Up to now, we have found that the word *dirty* can be labelled "core" with reference to Features 1, 2 and 3. But this is not the case with Feature 4. Looking up the members of the English lexical set comprising *dirty, filthy, grimy,* and *grubby* in *Roget's New Pocket Thesaurus* (Lewis, 1969) we find that *dirty* connotes "uncleanness", "meanness", and "obscenity", *filthy* connotes "uncleanness" and "obscenity", *grimy* connotes "uncleanness" and *grubby* "untidiness". As the carefull reader will see, core items do not have any connotations. This is at variance with the case we had with *dirty*, because it has three connotations. In other words, considering Feature 4 in Carter's model, *dirty* is labelled a "noncore item" as compared with *filthy, grimy,* and *grubby*. Therefore, it seems to be much better to assign fractional values for the degrees of coreness regarding this feature. Thus we obtain

grubby grimy	>	filthy	>	dirty (The non-core item considering feature 4)
3/4 untideness (only one connotation as determined by the thesaurus)		1/2 uncleanness obscenity		¼ uncleanness meanness obscenity

Using Noi'n's *Persian-Persian Dictionary* (Noi'n, 1983), the same analogy can be extended to the corresponding Persian lexical set *kasīf, nā-pāk, ālūdeh*, and *cherkin*. According to this dictionary, *nā-pāk* carries the most connotation. It connotes "obscenity", "untidiness", "religious-wise uncleanness", and also "meanness". *Kasīf* connotes "meanness" and "uncleanness". *Ālūdeh* and *cherk-in* connote only uncleanness. Therefore we obtain :

ālūdeh/cherkin	>	*kasīf*	>	*nā-pāk*
3/4		1/2 (the noncore item considering Feature 4)		0

- The application of Features 5 and 6

See Figure 4 for a very clear introductory account. For the sake of brevity we can say that, considering F5, all the items are labelled "core item". This can be accounted for by the fact one cannot say for certain to what particular field of discourse they belong. Therefore, the relative degree of coreness for all these items corresponds to 1 which is to be inserted in Figure 4. In addition, Feature 6 does not apply to our case, since we do not deal with nouns here.

	Relative Degrees of Coreness English				Relative degrees of Coreness Persian			
Carter's 6-item Model (core words are characterized by ...)	Dirty	Filthy	Grimy	Grubby	kasīf	nāpāk	ālūdeh	cherkin
Feature I (clear antonyms)	1	0	0	0	1	0	0	0
Feature 2 (high collocational span)	1	.53	.46	.23	1	.3	.23	.15
Feature 3 (are used to define...)	1	0	0	0	1	0	0	0
Feature 4 (No marked connotations)	1/4	1/2	¾	3/4	1/2	0	3/4	3/4
Feature 5 (no field of discourse)	1	1	1	1	1	1	1	1
Feature 6 (superordinates)	–	–	–	–	–	–	–	–
TOTAL DEGREES OF CORENESS	4.25	2	2.2	1.98	4.5	1.3	1.98	1.9

Figure 4 : Quantitative Analysis

(Quantitative analysis : *Dirty* was labelled "core" in terms of Features 1,2,3 and 4 in Carter's six item model, that is 4 out of 6. In the case of *filthy, grimy, grubby, kasīf, nā-pāk, ālūdeh*, and *cherkin* the numbers run as (2,4,5),(2,4,5),(2,4,5),(1,2,3,4,5)(2,5),(2,4,5),(2,4,5), respectively.)

This shows that the English core items may be translated by their Persian core counterparts. Ideally speaking, the same must hold true for noncore items; that is noncore lexical items in English must, theoretically speaking, be translated by noncore lexical items in Persian to observe the "dynamic equivalence" and the dynamic effect advocated by Nida (1982).

It was the writer's hope to show that noncore items can be, realistically speaking, translated by noncore items and vice versa. But, at least, this did not happen with Persian and English adjectives and the afore-mentioned statement proved to be far unrealistic except for a few cases where the noncore words were not too culture-bound to be handled. Instead, it was found that core items whose part of speech is adjectival may be fruitfully used for the noncore items with minimal risk of losing connotations and the collocational harmony. It is quite evident, however that no problems arise when one translates core items into their core counterparts in another language. For instance, *dirty* with the highest degree of coreness corresponding to 4.25 can be fruitfully translated into Persian using the word *kasīf* with a degree of coreness equal to 4.5 out of 6 (Every feature was assigned a maximum value of 1. Therefore, the highest degree of coreness corresponds to 6). More than that, these two can replace the other members of the two sets.

Taking a deep look at the degrees of coreness, one expects, for instance, *ālūdeh* having a degree of 1.98 to be a good equivalent for *grubby*, whose degree of coreness is also 1.98. Unfortunately this is not the case and numerous collocational and connotation-wise problems arise when one tries to do so. Therefore, you must see if the two words really collocate in the target language prior to putting them together. Consider the following:

Grubby hands (*ālūdeh* is a good equivalent)
Grubby scandal (This is at variance with the hypothesis which states that noncore items must be translated by noncore items. In Persian one cannot use *ālūdeh* meaning unclean with *rosvaie* (scandal). They do not normally collocate.
Grubby clothes (Appropriate)
Grubby face (*Ālūdeh* is not a good equivalent. It does not collocate with *sourat* (face) in Persian.)

Though *ālūdeh* can be used fruitfully as a good equivalent in 2 out of 4 cases discussed above, a simple 2/4 coverage does not seem to be a good generalization. But the same *kasīf* can be used together with all the 5 cases to be sure of keeping the connotations and collocations satisfactorily constant. Therefore, the core item *kasīf* can fruitfully replace the noncore *ālūdeh* to preserve the connotations and collocations implied.

Conclusion

In sum, the present paper makes an attempt to find a lexical equivalence for English adjectives with specific reference to the notion of coreness / noncoreness advocated by Carter. This being the case, the relative degrees of coreness were obtained, using Carter's six-feature model; the items having the most degrees of coreness were identified. Two lexical sets, one coming from Persian and the other coming from English, were taken as examples. In the end, it was found that, most noncore adjectives in English cannot be, realistically speaking, translated into noncore items in Persian. This can be attributed to the problems which usually accompany collocation, connotation, and the culturebound network. Instead, it was found that core adjectives in the source language can be translated by core adjectives in the target language. More than that, the same core adjectives can, most often, serve as best equivalents in the absence of appropriate noncore adjectives in English. For instance, the core adjective *kasīf* can almost always replace *ālūdeh* in Persian especially when problems with connotation and collocation arise. For sure, the identification of core words in the two languages can help the translator in this context. This conclusion, however, cannot be carried too far in absolute terms. In other words, it can only be limited to the translation of adjectives.

References

CARTER, R. 1988. "Front Pages : Lexis, Style and Newspaper Reports" in *Registers of Written English : Situational Factors and Linguistic Features,* Ed. by Ghadessy, M., London : Printer Publishers, 8-16.

CARTER, R. 1982. "A Note On Core Vocabulary", Nottingham Linguistic Circular, Nottingham, 39-50.

CELCE-MURCIA, M. and McINTOSH, L. 1979. *Teaching English as a Second or Foreign Language.* Cambridge : Newbury House Publishers (p. 242-256).

ELLIS, R. 1985. *Understanding Second Language Acquisition*, Oxford: Oxford University Press (p. 190-194).

HORNBY, A.S. 1989. *Oxford Advanced Learner's Dictionary of Current English.* Oxford : Oxford University Press (p. 454, 548, 553).

LEWIS, N., 1969. *Roget's New Pocket Thesaurus,* New York : Washington Square Press. (p. 114,159,183,184).

NIDA, E.A. & TABER, C.R. 1982. *The Theory and Practice of Translation,* Leiden : Brill (p.22-23).

NOI'N, N. 1983. *Farhang-o-Farsi,* Tehran : Entesharat-e-Amir Kabir (p.84,1282,2909-2910,4526).

RICHARDS, J., PLATT, J., WEBER, H., 1989. *Longman Dictionary of Applied Linguistics,* Hong Kong : Longman (p. 171).

About the author

As a BA student, Abdol Mohammad Jahangiri showed great interest in etymology thus giving expression to a new-method fast-find mini-dictionary covering some interlingual words which draws upon a contrastive analysis of English and the other languages of similar families (French, English, Greek, Latin, Old Persian, etc.) to find the missing links between

the languages. Up to the end of the BA programme he managed to establish the scientific errata of several dictionaries including "Oxford Advanced Learners' Dictionary of Current English", Merriam Webster", "Collins English Paperback Dictionary", etc. He received his MA from the Azad University of Shiraz. Among his current research interests are Old Persian, text analysis, translation theory, F.S.P. and translation, markedness theory, Core/Noncore vocabulary and translation, politeness forms, the translation of technical texts, etc.

Address : Abdol Mohammad Jahangari, Saheli-e-Sharghi, 8metri Avval,
107, Post code 71457, SHIRAZ, IRAN, Tel. 071-26539

Abstract

Objective measurement in translation evaluation is a relatively untouched area, especially at the level of the lexis. No thoroughly worked-out method has been put forward for the scientific evaluation of a translation. Yet few translators would question the potential value of such a device.

Certainly translation evaluation is not limited to the lexis, but it is the vocabulary items which form the building blocks that make up the whole text. So perhaps this is the best level for translation evaluation to start at, that is lexical translation evaluation. The article focuses on how determination of the relative degree of core-ness of a given vocabulary item (i.e. the degree to which it belongs to the core vocabulary of the language) can help one find the best equivalent in another language (hence COLTE). Adopting a quantitative approach, we attempt to devise a scalar test of lexical equivalence.

Résumé

L'évaluation des traductions est un domaine peu abordé, surtout au niveau lexicologique. Il n'existe à la connaissance de l'auteur aucune méthode scientifique pour évaluer une traduction. Or aucun traducteur ne contesterait la valeur d'un tel outil. Bien entendu, l'évaluation d'une traduction ne se limite pas au seul aspect lexicologique. Ce sont toutefois les aspects liés au vocabulaire qui créent le texte. Sans mots, il n'y a pas de communication. En soi, cet énoncé est une bonne base de départ pour évaluer une traduction du point de vue lexicologique. Dans son article, l'auteur examine dans quelle mesure l'identification du degré de centralisation relatif d'un élément lexicologique donné peut aider le traducteur à trouver le meilleur équivalent dans la langue d'arrivée (ou ELTOC). Le présent article adopte une approche mathématique (ou plutôt quantitative) pour présenter une échelle de fidélité lexicologique.