

Notation and Index for Informative Congruence in Translation

Brian Harris

Volume 20, numéro 3, septembre 1975

URI : <https://id.erudit.org/iderudit/003031ar>

DOI : <https://doi.org/10.7202/003031ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Harris, B. (1975). Notation and Index for Informative Congruence in Translation. *Meta*, 20(3), 184–193. <https://doi.org/10.7202/003031ar>

Notation and Index for Informative Congruence in Translation¹

I. PURPOSE

The idea of somehow « measuring » translation may seem surprising but it is far from new. It was adumbrated for example by I. A. Richards, who also pointed out that any measurement requires a prior selection of what « qualities » are to be measured :

When measuring a room we give measurements *in respect of* height, length, etc. So in measuring translation we need a system of « respects ».

The easiest thing in respect of which to measure any text is its surface bulk, i.e. its length, most conveniently recorded as the number of words. Van Hoof has ventured an evaluative criterion based directly on this quantitative parameter : « *À valeur communicationnelle égale, la traduction la plus courte sera la plus pertinente.* » But if we accept, along with most writers on translatology including Van Hoof by virtue of his « *valeur communicationnelle égale* », that the main purpose of most translations is to re-code a *message* (« message » in the sense of « some information »), then what we could like to be able to measure by hook or by crook — by which I mean by indirect methods if it cannot be done directly — is the preservation of information and conversely the garbling or complete loss of it. In short the prime object of our study will be *content* ; the study of expression will only concern us insofar as it is a means to that end, a source of outward « symptoms » of the inner fluctuations².

1. The term « congruence » is taken from Laurent Tarnoczi's article : « *Congruence entre l'original et la traduction* », *Meta* 13:3 (1967) 137-143, where he says it means the same as Fedorov's « *adekvatnost* », Jumpelt's « *Aequivalenz* » and Nida's « *conformance* ». To these terminological variants can be added Kade's « *invariance of content* » (O. Kade, *Zufall u. Gesetzmaessigkeit in der Uebersetzung*, 1968). Tarnoczi begins by criticizing the older concept of « *fidelity* » : « *notion vague que personne jusqu'ici n'avait suffisamment concrétisée afin de la rendre utilisable pour la qualification de la traduction.* » He goes on : « *Seule, la confrontation des deux textes en présence permet d'aboutir à un résultat concret... Une traduction et un original peuvent être considérés comme congruents s'ils expriment la même teneur informative, la même intention communicative et ont sur les lecteurs ou auditeurs le même effet.* »
2. The Richards quotation is from « *Towards a theory of translation* » in A. F. Wright (ed), *Studies in Chinese Thought*, Chicago, 1953. For Van Hoof, see the next Note. « *Translatology* » is the term I proposed in a 1972 lecture as more concise and more scientific-sounding than « *theory of translation* » (see « *La traductologie, la traduction naturelle, la traduction automatique et la sémantique* », *Cahiers de linguistique de l'Université du Québec à Montréal*, n° 3, 1973, p. 133-146).

If we are to do any exact research on translation — heeding Ljudskanov's advice to welcome « the introduction of exact research methods into the social sciences... amongst which is translation theory » — we shall need exact notations in which to record our observations. The notation described below has been devised for use in one method of translation research, namely the comparison of translated texts with their originals and with one another³. It has already been modified several times in the past year in response to shortcomings uncovered by students' attempts to apply it. Apart from research, it may be useful to teachers of translation, though it must be borne in mind that it does only annotate content and not expression⁴. This is not to imply expression is unimportant in translation : only that although content and expression are interwoven, it is possible and profitable to analyze them separately⁵.

By « content » I mean all the information that the author of the original text apparently intended to convey. « Apparently » is a qualification which is forced on us in all but the exceptional cases where the author can be consulted in person or through an exegesis in his other writings. « Intended » is a constraint which implies conversely that a text may reveal more than its author realizes ; and then the translator faces a choice between passing on this revelation or suppressing it. A classic question of apparent intent arises when an SL text contains dialect or archaisms. So far as we know his mind, is the author using them unwittingly because that is his usual way of speaking, or have they been put in for « local colour » ? If the latter, then the use is intentional ; and if the geographical or temporal location they signal is not signalled in the translation in some way or other, then there will be loss of information.

II. SEGMENTATION

The assumption is made here that the texts, both source language (SL) and target language (TL), have first been segmented into « translation units ». Van Hoof defines a translation unit as « *le plus petit assemblage de mots qui contribuent à l'expression d'un seul fragment de message et dont le degré d'interdépendance est*

3. The beginnings of such a notation are found in H. Van Hoof's 1971 article « Recherche d'un modèle d'analyse en traduction » (*Meta* 16 : 1-2, 83-94), and no doubt others have tried. But Van Hoof provides only three symbols, one of which (M = *traduction la plus pertinente en fonction du message*) is a catch-all.

The Ljudskanov quotation is from his *Mensch u. Maschine als Uebersetzer*, Halle (Saale), 1972 : « Die Einfuehrung von exakten Untersuchungsmethoden... in den gesellschaftswissenschaftlichen Disziplinen — darunter in... der Uebersetzungstheorie » (p. 12).

4. Of course we only have access to the content through the expression, and it is on the written substance of the latter that we have to append the notation. « Translation units » are segments of *expression*, but they are intended to correspond to units of *meaning* that have a certain viability for purposes of translation. This two sidedness of the translation unit comes out in D. Goudec, *Comprendre et traduire*, Paris, 1974, where he proclaims, « *nos unités seront des unités syntaxiques* », and then a few lines later : « *L'unité de traduction... c'est la plus petite partie de l'énoncé qui suffit à évoquer un élément complet de la situation décrite.* »

5. « *L'harmonie de la forme et du contenu ne doit pas être interprétée uniquement comme si le contenu ne pouvait être exprimé que dans une seule forme nécessairement donnée.* » — Tarnoczi, *op. cit.* In other words, we are attempting to apply only some of the components of the kind of « *modèle global* » that Van Hoof aims at in *op. cit.*

*tel qu'ils ne peuvent se traduire isolément dans le cadre du message total*⁶. However, though we think we know what he means, this informal kind of definition does not ensure that all researchers given the same text will choose the same unit boundaries. For a procedure that produces fairly consistent segmentation, see the *Appendix* below. Either a vertical line or a slant may be used to mark the boundaries between units, and double lines for sentence boundaries. The symbols provided in what follows should be written over the units of the TL text, to the right of the unit number. An example will be given in the next section.

III. BASIC SYMBOLS

- = No change.
- ∅ Entire loss of a SL unit. Since these units do not appear in the TL text, the SL unit number and the symbol ∅ should be written over the boundary line between the TL units that correspond to the SL units bordering the lost one.
- Partial loss of a SL unit.
- ⊕ Translator's invention of an entire TL unit. These added units are to be numbered na, nb,... where n is the immediately preceding unit number.
- + Partial addition in a TL unit.
- n ∈ n' (where n and n' are unit numbers) Fusion of the two correspondingly numbered SL units into one TL unit.
- ≠ Some other change.

Example 1 :

SL	1	2	3 4
	« //Ça fait déjà plusieurs jours/qu'il passe plus/par ici/?// »		
TL	2+ ∈ 3	1—	1a+ 4=
	« // Tom hasn't been around/ for several days/ I hear/ ./ / ⁷ »		

Example 2 (from an East German publication) :

SL	1	2
	« /wie man sie auch nennt,/von kybernetischen Maschinen/ »	
TL	2	1 1a
	« /by cybernetic machines,/as they are also called/in the USSR/ »	

6. *Op. cit.*, p. 89. The condition « *le plus petit* » turns out to be relative, and so later in his article Van Hoof introduces a dichotomy between « *grosses unités* » and « *petites unités* ».

7. Unit 4 is marked « = » to illustrate a point. The change of expression from a period to a question-mark does not necessarily constitute a change of content. We have also to consider what Tarnoczi calls « *communicative intention* ». Is the question form truly intended to draw forth information in response, or is it just « rhetorical » ? If the former, they by all means let us annotate 4≠ ; but if the latter, then 4= is proper.

J The author's judgement of circumstances¹¹. Cf (from two translations of *Othello*) : « *Je serai son lieutenant* » vs « *Je serai forcément son lieutenant* ».

L Logical connections.

Ex : 1 2
 SL « /He liked it/so be bought it// »
 2 ≠L 1
 TL « /Parce que ce morceau lui plaisait/il l'acheta// »

M Metaphor¹². This is the most disputable of the indicators, since it can be argued that metaphor is a matter of expression, not content. However, we would make the distinction between live metaphors and dead ones. The latter are of no interest for our present purpose : for instance there is no need to translate « river mouth » by « *les bouches du fleuve* » rather than « *l'estuaire* ». The argument for including *live* metaphors on the other hand is that they add information : this is based on the idea that metaphors are derived from similes and that similes make comparisons, e.g. « An opening into Hell » + « The opening looks like a pair of gaping jaws » => « An opening like jaws into Hell » => « The jaws of Hell ».

—M The reverse of +M, i.e. replacement of a live metaphor by a more literal expression.

Ex : 1
 SL « / He shed crocodile tears/ »
 1—M—C
 TL « /Il s'est montré hypocrite/¹³ »

P Precision.

Ex : 1 1 +P
 SL « /early in the morning/ » TL « /à l'aube/ »

—P The converse of +P, i.e. loss of precision, more generality.

Ex : The reverse of the last example.

R Respect. Cf. F above.

11. J corresponds to what T. E. Alisova, « Dopoltinnye otosheniya modusa i diktuma », *Voprosy Yazikoznaniya*, 1 (1971), 54-64, calls the « modus », which she defines as the speaker's attitude towards the subject-predicate relations of a denotation (« dictum »). The predicative nature of each utterance is composed of these two components.

12. Van Hoof, *op. cit.*, uses the symbol I (for « *équivalence imagée* »).

13. —C because the French unit fails to inform that the hypocrisy is in a show of grief. Note that I speak of a (relatively) « more literal » expression. The boundary between literals and metaphors is fuzzy, for one thing because of the « dead » metaphors. Furthermore « dead » metaphors may be more or less dead, dead for some people yet not for others. Even in the fragmentary French example given here as « more literal » we can spot two metaphors : « *se montrer* » and « *hypocrite* ». Are they dead or alive ? Surely dead for all but the very learned : but cf. *Petit Larousse sub voc.* « *montrer* » and « *hypocrisie* ».

U Cultural setting.

Ex : (from *Hier les enfants dansaient*) :

1 1 —U

SL /« *Bonsoir !* »/ TL /« *Ah !* »/¹⁴

W Weakening or detensification, the converse of I.

Ex : 1 1 —I

SL « /*bête*/ » TL « /rather silly/ »

X Explicit statement of implications and presuppositions. The implication must be *entailed* by the text : explanatory material added by the translator on his own initiative should be marked ⊕.

Ex : (from an East German publication) :

1 2

SL « / *fuehrte/zum Formalismus/* »

1 2 +X

TL « /would lead/to « formalism » in its pejorative sense/¹⁵ »

Z *Implied* information about the author/speaker himself, other than what can be represented by the symbols already given. (If the information is stated explicitly, it is to be treated as belonging to the cognitive content.) This covers what linguists call « illocutionary meaning » or « performative verbs ». An example would be the change from an affirmative sentence to a question ; because the former implies that the speaker knows something, whereas the latter implies there is some knowledge he is lacking.

? A doubtful annotation. It must be placed immediately before the symbol to which it applies, e.g. —?E means there is definitely a loss but the investigator is not certain whether historical period is affected.

! Draws special attention.

* Before a symbol, it may be used to mark a wrong annotation. Note, however, that it is the *annotation* which is wrong in such cases, not the translation.

V. INDEX OF INFORMATIVE CONGRUENCE

The purpose of the Index is to provide a highly condensed measure of congruence in a translation for the purpose of comparison with *another translation of the same text*.

14. « *Bonsoir* » is a polite salutation required by French good manners. « *Ah !* » is offhand. Yet this is the opening line of a play about French-Canadian society.

15. The translator illuminates for Western readers the aspect that the word takes on in its Marxist context.

This index differs in two ways from earlier methods of quantifying translation¹⁶:

- (i) It is based on a count by « translation units » instead of by number or categories of *words*.
- (ii) It takes into account *partial* changes as well as complete loss; and it also penalizes for *additions* to the original message.

On the other hand, at its present stage of development it is based solely on analysis of the segments, which constitutes a sub-sentence level. Further weighting ought to be built in to take account of sentence and whole-text levels. Moreover it can be objected that there are only two weights, 1 or ½, and that this is too crude a scale.

Be that as it may, the formula as it stands at present is:

$$I = \frac{a + \frac{1}{2}(b + c + d) + (0 \times e) \times 100}{s + f - g}$$

where I is the index figure

<i>a</i>	is the number of occurrences of	=
<i>b</i>	” ” ” ” ” ”	—
<i>c</i>	” ” ” ” ” ”	+
<i>d</i>	” ” ” ” ” ”	≠
<i>e</i>	” ” ” ” ” ”	∅
<i>s</i>	is the number of units in the SL text	
<i>f</i>	” ” ” ” occurrences of	⊕
<i>g</i>	” ” ” ” ” ”	∈

The reason for multiplying by 100 is to raise the index figure to a number that can be rounded off to an integer. If there is more than one of the symbols —, +, ≠ on the same segment, only one of them is counted. It is useful for checking to bear in mind that (*a* + *b* + *c* + *d* + *f*) equals the number of TL units. (0 × *e*) is of course a constant zero and is only mentioned in the formula to show what happens to the ∅ units in the calculation. *f* is added to *s* so as to lower the index if the translator adds « inventions » of his own. *g* is subtracted because each « ∈ » implies that there is one SL unit which is missing from the *form* of the TL but which ought not to affect the index by its elision since its *content* has been retained.

Ex: SL text has 250 segments. TL text has 145 segments marked =, 48 segments marked —, 30 segments marked +, 5 segments marked ≠, 10 segments marked with one ∈, 2 segments marked with two ∈s, and 10 segments marked ⊕. ∅ does not occur.

$$I = \frac{145 + \frac{1}{2}(48 + 30 + 5) + (0 \times 0) \times 100}{250 + 10 - 12} = \frac{18650}{428} = 75.2$$

16. Notably Gilbert Barth, *Recherches sur la fréquence et la valeur des parties du discours en français, en anglais et en espagnol*, Paris, 1962, summarized in part as « French to English: Some stylistic considerations », *Meta* 16 (1971), 33-44.

Ultimately the value of the index depends on its correlation with the reader's *understanding* of the translation either in whole in or part — which would have to be determined by unilingual comprehension tests of the SL text — and with other criteria of reader acceptance.

VI. EXPERIMENTAL METHODS

The main problem in the scientific application of the notation is that the descriptions accompanying the symbols are informal and therefore subject to just the same criticism as I levelled earlier against Van Hoof, namely that they do not ensure different investigators will mark the same text in the same way. Notwithstanding much recent progress in semantics, including the development of computational semantic analysis¹⁷, it is not yet feasible to make a thoroughly consistent semantic analysis of substantial lengths of text nor one with all the detail required for the auxiliary symbols.

For the moment three compromises seem possible to me. One is the expedient of having several investigators do each analysis as a team and arrive at a consensus. It is what we have been doing so far.

The second is to provide investigators, individually or as a team, with some operational criteria for their decisions. This might be done by drawing up questionnaires to be answered for each segment in each text. The questions might be on the lines of : Does the author appear to approve of what he is describing ? Or disapprove ? In what historical period would you situate the things described ? And so on. Afterwards the answers from the examinations of SL and TL texts would be compared. A variant of this procedure would be to have the questions answered by native-speaker informants, one for SL and one for TL, instead of by the investigators. Obviously this variant would be more objective though longer.

A third possibility is to make up special SL texts and have them translated. In this way the major initial problem would be settled of determining what the SL author intended. It is quite a different situation, of course, from annotating existing translations. For experimental purposes, however, it affords the best approach because it should enable the investigators to control the content that is being translated. It might even be possible to devise behavioural tests such as composing a text describing an object, having it translated, and then asking two independent informants to pick out the object, one of them after reading the SL text and the other after reading the TL.

VII. CAVEAT

It may be as well to emphasize, in conclusion, that neither the symbols nor the index are intended to distinguish between « good » and « bad » translations.

17. To mention but a few, Brown's BELIEVER system at Rutgers University, Chafe's Contrastive Semantics Project at Berkeley, Schank's Conceptual Dependency system at Yale. For some idea of the current enthusiasm, see the collection of abstracts in *Linguistic Society of America Thirty-Sixth Summer Meeting & Association for Computational Linguistics Twelfth Annual Meeting: Meeting Handbook*, Secretariat of the LSA, 1974.

Value judgements may be based on them, but in themselves they should only be used to annotate *convergence* and *divergence* of information content as objectively as possible.

In any event it is a platitude to admit the impossibility of conserving everything from the SL text in a translation. We may end, as we began, with the thoughts of I. A. Richards : The translator's aim must be minimal loss. « What sort of loss will we take in order not to lose what ? »

* * *

APPENDIX

One major difference between the linguist and the translator is that the former can choose to analyze syntax and lexicon separately, whereas the translator must take both into account simultaneously. Therefore segmenting a text into « translation units » requires both syntactic and lexical criteria, and we have insufficient reason so far for supposing that either kind is of prior importance¹⁸.

The guidelines that follow are for use in studies comparing a SL text with its TL equivalent. Whether translators in the act of translating actually proceed consciously or intuitively by these criteria, is a proper subject for investigation by future students in course TRA 4901/3955 using methods like those we have employed in the « Schoolboy Latin » study.

GUIDELINES

1. Punctuation

Respect the punctuation, which should, so far as it goes, accord with both the lexical and the syntactic criteria given below. It provides some rough guide to the author's own segmentation.

— End-of-sentence punctuation marks, i.e. period, question mark or exclamation mark, are to be marked as separate segments because they give *information* about the author/speaker. An affirmative sentence, unless it explicitly states the contrary, implies that the author knows something ; a question usually means that he does not, and an exclamation means that he is very surprised and feels strongly.

2. Lexical criteria

This takes into account « idioms » and more.

— Do not split compound lexemes, e.g. English two-word verbs.

— Take together all lexical « collocations », e.g. those that make up Lacroix's *Dictionnaire des mots et des idées*, on the phrases given in Harrap's¹⁹.

18. In our operational study of high-school students translating Latin, though, we have noticed a marked tendency to search first for the verb. The report and data from this study are available in the Linguistics Documentation Centre. It is interesting that the verb-first strategy is predictable from Tesnière's dependency grammar (*Éléments de syntaxe structurale*, Paris, 1959) provided we assume « top down » analysis of sentences. It also agrees with Goudec's instruction in *op. cit.* : « *Le repérage de l'unité de traduction... se fait par repérage du verbe.* »

19. U. Lacroix, *Dictionnaire des mots et des idées*, Paris, 1963 ; J. E. Mansion (ed), *Harrap's Standard English and French Dictionary*, London, 1934-1939.

3. *Syntactic criteria*

Take as units *simple* phrases (NP, VP, etc.)

— Split compound and complex phrases into constituent phrases., separating off embedded clauses.

— If there are several adjectives to a noun (or adverbs to a verb) take the head word together with one of the epithets and treat the other epithets separately one by one²⁰.

4. *Enclitics*

Enclitics form part of the simple phrases envisaged in guideline 3, so should not be split off. « Enclitics » is used here in an unusually broad sense to include demonstratives, articles, personal and possessive pronouns and quasi-pronouns, auxiliary verbs, numbers and other quantifiers, and even sentential demonstratives like « *Il y a* » and « *It's* ».

5. *Prepositions*

The general rule is that prepositions go with the NP or VP that they govern. The major exception is the « *préposition fortement régie* » : these are prepositions required and selected by a preceding noun, verb, adjective or adverb, e.g. « different from », « *donner q.c. à q.u.* ». Most of them are specified in the big dictionaries like Harrap's.

Note : This rule does not mean one *must* start a new unit at a preposition. If one did, it would lead to absurdities like « *pomme/de terre* », contravening guideline 2.

6. *Conjunctions*

— Group *coordinating* conjunctions with the first unit of the coordinate that follows the conjunction.

— Likewise group *subordinating* conjunctions, for instance complementizers and relative pronouns, with the first unit of the subordinate clause. e.g. « The man / who came ».

iii. Leave *inter-sentential* connectors (« *charnières* ») as separate units, e.g. « *Néanmoins / il me semblait...* ».

7. *Adverbs*

— For *verb modifiers*, follow the same guideline as for adjectives (see guideline 3)

— *Adjective modifiers* go in the same unit as the adjective

— *Sentential adverbs* are separate units.

Examples (from Saint-John Perse's *Anabase*)

1. « *J'augure bien d/u sol* » (guidelines 2,4,5. Harrap's gives « *augurer bien de q.c.* »)

2. « *Gens de poussière / et de toutes façons* » (guidelines 3,5,6)

3. « *... / tient ses assises/* » (guideline 2)

BRIAN HARRIS

20. Collocations are an important part of the lexicon, e.g. the fact that one says « I am *badly* in need of s.t. » but « I am *madly* in love with s.o. ». Monolingual dictionaries hardly deal at all with collocations, though Lacroix is an honourable exception. The good bilingual dictionaries do better but not nearly enough. The current work in the USSR on « semantic parameters » is a tentative attempt to fill this need methodically : see Y. D. Apresyan, I. A. Melchuk, A. K. Zholkovsky, « Semantics and lexicography : Towards a new type of unilingual dictionary » in F. Kiefer (ed.), *Studies in Syntax and Semantics*, Dordrecht, 1970.