

A Proposal for Dictionarization of an Indian Language

Marília Facó Soares

Volume 41, Number 2, juin 1996

Traduction et terminologie au Brésil
Translation and Terminology in Brazil

URI: <https://id.erudit.org/iderudit/004198ar>
DOI: <https://doi.org/10.7202/004198ar>

[See table of contents](#)

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (print)
1492-1421 (digital)

[Explore this journal](#)

Cite this article

Soares, M. F. (1996). A Proposal for Dictionarization of an Indian Language.
Meta, 41(2), 288–294. <https://doi.org/10.7202/004198ar>

Article abstract

This article presents a proposal for a dictionary whose preparation relies partly on Indian knowledge. The Indians themselves would be active participants in this process for which the author has coined a term: dictionarization. The article concludes with a correlation between dictionarization and the translation of conceptual terms from an Indian language.

A PROPOSAL FOR DICTIONARIZATION OF AN INDIAN LANGUAGE

MARÍLIA FACÓ SOARES

Department of Anthropology, National Museum/UFRJ,

Résumé

Cet article discute d'une proposition pour l'élaboration d'un dictionnaire auquel les Amérindiens pourraient participer activement. L'auteur appelle «dictionnarisation» une telle participation qui aurait pour effet de consacrer le dictionnaire aux yeux des Amérindiens. On discute du processus de «dictionnarisation» pour soutenir le projet de dictionnaire et on termine sur une corrélation entre notre proposition et la traduction de termes conceptuels dans la langue concernée, le tikuna.

Abstract

This article presents a proposal for a dictionary whose preparation relies partly on Indian knowledge. The Indians themselves would be active participants in this process for which the author has coined a term: dictionarization. The article concludes with a correlation between dictionarization and the translation of conceptual terms from an Indian language.

It is a well-known fact that anyone working with Indian languages will come across a number of difficulties in translating conceptual terms from those languages. In addition to the problems intrinsic in translation *per se*, a good many researchers encounter other problems since they also wish to provide feedback regarding indigenous languages to the Indians themselves. And this is often achieved through some form of dictionary — a material piece of work that ultimately involves translation. The key purpose is to help the Indians,¹ in addition to providing a tool to those who are in contact with indigenous peoples, whether out of professional need, intellectual curiosity and/or commitment to the Indian cause.

Not unlike our colleagues, at a certain point we felt that our efforts should result in something useful to the native groups who speak the language we were studying. Our goal was first and foremost not to help the Indians but rather to integrate the analytical possibilities afforded by our work and the analytical potential of the speakers of the language under scrutiny.

We believed that one consequence of this effort was that the Indians themselves would be part of the process. We even coined a term for this process: *dictionarization*. And this is the main issue discussed in this paper, as we develop a proposal to support such a process. In our conclusion, we draw a correlation between our proposal and the translation of conceptual terms from an Indian language.

A PROPOSAL FOR DICTIONARIZATION

The proposal we have developed focuses specifically on forms of the Tikuna language believed to be genetically isolated and spoken by a large population living outside the Brazilian borders (in Peru and Colombia), and around a vast expanse of the Upper Solimões area covering eight municipalities in the state of Amazonas (Tabatinga, Benjamin Constant, São Paulo de Olivença, Amaturá, Santo Antonio do Içá, Tonantins,

and Beruri). Over 20,000 Tikuna are spread out in 69 settlements. The Tikuna have been in contact with surrounding communities for a long time (roughly 300 years).

Let us now clarify the three key words used throughout this proposal: *lexicon*, *dictionary* and *dictionary*. The first of them concerns the internalized knowledge shared by speakers of the realm of words employed in their own language, *i.e.* knowledge involving the words of that language, their different meanings, their underlying regularity, different usages, possibilities of syntactic use, and correlations between and among words. The second term, *dictionary*, is directly connected to the existence of a material object which, depending on how it is constructed, may be used to reflect features of the lexicon. The last of those three terms, *dictionary*, is employed to cause a shift: the emphasis that would normally be on a dictionary as a material object, would cover the very process leading to the compilation of a dictionary. The process, the material object to be produced, and what it will reflect constitute a topic yet to be developed in linguistics projects dealing with Brazilian Indian languages.

The Tikuna have played a key role in the proposed dictionary, both during the survey and selection of sources, grouping and classification of words and expressions, and throughout the process of selecting entries and entry organization. The reason for this was that, in our view, the handling of a cultural product by those who will be their users very much depends on how involved the society is with the action giving rise to the product in question. In societies where knowledge is specialized, delegating the task of giving shape to something that not everyone is expert enough to do, and from which all will benefit does not usually affect the common feeling that what is done belongs to the entire group. In societies where knowledge is not compartmentalized or where such compartmentalization is just beginning to be felt with the introduction of a specialized body of knowledge² due to some historical situation, it is apparently not easy to combine that specialized knowledge and the shared feeling of acquisition of cultural products by the group, particularly when the knowledge at stake has a source external to the group. In the latter case, the resulting cultural asset remains an external item to those who, in principle, should use it. In terms of the existence of a dictionary within an Indian group — an external cultural item per se — we believe that its effective utilization will occur to the extent that integration is achieved between the new dictionary and characteristics of daily life among the target Indian group. This integration, however, will not be attained through any action that *talks of* something, but rather through action that *talks with* it. So the idea of a bilingual dictionary as a finished product is discarded; this would be a material good resulting from the efforts of a scholar who donates part of his or her organized knowledge to a certain Indian group. And since the desired integration is associated with the action of *talking with*, the focus here shifts from a product called dictionary to a dictionary conceived as the outcome of a process, where the process is as important as its end result. This is why we talk here of dictionary rather than of dictionary.

To bring this proposed dictionary to the realm of a specific indigenous group, the Tikuna, we must first of all determine the scope of the action that *talks with* while at the same time defining the possibilities for real participation of the target Indian population.

AGENTS IN THE DICTIONARIZATION PROCESS

As indicated, the actions described here involve more than one individual. Among those who might become involved in such actions are individuals already doing research with the Tikuna, others who are experts in certain fields of knowledge and would be very helpful, and the Tikuna themselves. These people can produce a collaborative effort.

Since the current project deals with the issue of languages (and here it does not matter whether it is one, two or more languages), and since the language brings a people's culture to life, it is in language that meanings become obvious, confront each other, intertwine, echo in words, expressions and text. If a dictionary, is conceived as place which embodies, in addition to the list of signs already developed, a reference to lexical regularities as well as correspondences, it can not leave out the related meaning³ between and among forms. Therefore people who are to become involved in its development cannot be expected to have their role limited to that of surveying and compiling an inventory of signs related to their own area of expertise only. What is expected of the persons involved is that, with their focus on texts in the Indian language, they may identify and group together forms to be subsequently defined in line with the meaning correlations they may have. A concrete example of how this type of participation might become effective is the work of linguists, in and of itself tied to the work of other individuals.

Linguists as such will not be the ones to establish criteria and preset categories which the data of a language must meet to fit into a theoretical model. They must apply modern linguistic principles derived from research into a good number of natural languages, including the so-called Indian languages. Among such principles are those that focus directly on development of a language lexicon:

- a) in a lexicon signs, in the Saussurean sense of the word, and morphemes, as the minimal meaningful unit, must not be equated with the sign since natural languages have morphemes with no related significance; signs in natural languages tend toward the word;
- b) a lexicon has standard and consolidated formations, with several stages of non-regularity coexisting, that is, several stages that stand between complete regularity and a total separation between a word's meaning or usage and the meaning or usage of its morphological parts;
- c) a lexicon has a number of naturally occurring metaphors and meaning extensions;
- d) a lexicon plays a key role in understanding a language's grammatical relationships in as much as lexical items, associated with semantic roles, require a certain structure in arguments; lexical items have features that allow a distinction to be made within a group of items forming a predicate and within a group of items related to the attribution of a syntactic role.

Adopting the above principles means departing from the linguist's monologue, and introducing other viewpoints between linguists and their data. Even a superficial examination of these above principles can help show that dialogue is required.

Let us start with the first assertion that in a lexicon, the sign tends toward the word. It is a well-known fact that in languages without a written tradition, words based on a single criterion cannot be identified. There are several criteria: understanding a word as a unit of meaning, the use of potential pauses, the concept that words are the smallest units in a language that may be used alone as a sentence, identification of the word in terms of minimum positional mobility unit, the concept that words are units with an internal fixed structure, that is, that they have internal stability. Which of these criteria should be selected or combined is a problem that the linguist has always had to face every time he or she is confronted solely with oral reality. Since words are easily recognizable despite the difficulty in defining them, if a linguist wants to work with words whose definitions are not the subject of a simplifying approach,⁴ the analysis will have to be made not from the standpoint of unilateral identification, but from that of recognition. Recognizing words is, understandably, easy for native speakers: they are the basic linguistic units, and recognizing them, even without recourse to writing, constitutes no problem for those who deal with them intuitively. The recognition of units through intuition enlists the aid of yet another player. Between linguist and data, comes the native speaker. And the compiling of a

dictionary — a cultural good in which words are handled. The perpetual issue as to what a word is, the distinction between words and the forms they take must necessarily require more than one linguistic standpoint.

The second issue has to do with the consolidated forms and the standard forms, with the existing stages of non-regularity in a language's lexicon. Minor sub-standardizations, minor deviations or even full irregularity in terms of connection between meaning or usage of a word and the meaning or usage of its component parts also require more than one linguistic standpoint. However, here the issue is more sophisticated. With words that are made up of a number of elements and show a semantic change without any morphological alteration of their parts, the minor deviation or total lack of connection between the overall meaning of those words and their morphological characterization requires a lexicon to include words of differing backgrounds,⁵ *i.e.*, the lexicon must contain words that semantically and morphologically indicate what has happened linguistically just as there are words that, because of their regularity, indicate what is and may still be happening from a linguistic standpoint. The lexicon of a language is therefore the place where part of its linguistic history is stored. At the same time, it is the place where one finds the linguistic history and the changes that have affected or may be affecting speakers of that language: the semantic evolution in words that leads to a gap between overall meaning and morphological characterization is commonly related to changes — incidental or otherwise — taking place outside that language but affecting it nonetheless. What the lexicon of a language tells us about consolidated forms and regular forms is relevant to the compiling of a dictionary. A dictionary may contain and list forms belonging to different stages of non-standardization, since the organization of entries will involve not only a definition of words but also their morphological breakdown. The degree of connection between overall meaning and partial meaning can be inferred therefrom, and with it the degree of irregularity the lexicon of a given language contains as well as the degree to which it mixes factors both of a linguistic and nonlinguistic nature. However, the effort to include in a dictionary what is taking place in the lexicon regarding non-regularity stages, requires more than just the linguist and native speaker. Because of the necessity to deal with linguistics and what lies beyond, one must consider the living standards the native speakers of the language to be consigned to a dictionary. Such conditions, involving both time and space, cover areas of activity and social interaction. Detecting and pinpointing meanings and special usages of words, recognizing metaphors and polysemy require input from people other than linguists.

Metaphor and extension of meaning have already been listed (together with narrowing of meaning, metonymy, synecdoche, hyperbole, litotes, degeneration and elevation of meaning), as well as classes of changes in meaning.⁶ However, when one looks at linguistic manifestations from a synchronous standpoint, there is no way of maintaining one-way meaning realms. Synchronically speaking, words may be considered to have a central meaning to which a whole range of meanings can be added. Among these are those relating to the core meaning through metaphors or extended meanings (or through metonymy, synecdoche, etc.). Metaphors and extended meanings, like other sense relationships, are not just semantic alterations introduced to the lexicon along a time scale (diachronic axis). They are semantic possibilities present at a particular moment in a language and that can be explained through culturally inserted conceptual models.

When it comes to compiling a dictionary, the fact that sense relations exist in the lexicon of a language, such as those created by metaphors and extensions of meaning, emphasizes the need to obtain assistance of persons familiar with the cultural universe of the speakers of that language one is attempting to consign in a dictionary. After all, sense relationships do not exist in a void (they are culturally inserted) nor do they depend on an

assumed mechanism operating within the language alone. Thus, if a dictionary is to reflect the characteristics of a lexicon, consideration must be given primarily to a dialogue that will involve not only the linguist and the native speaker, but also all other persons who are familiar with a given cultural universe.

We have mentioned grammatical correlations as a fourth issue related to the lexicon. This would appear to put us in a purely linguistic field where, for representation of grammatical features in the lexicon, it would suffice to have a linguist with his or her set of data. However, the very treatment given to grammatical correlations is no longer entirely discrete. Even if one were to adopt a purportedly independent line of syntax,⁷ aspects relevant to syntax are so dependent on the postulation of semantic roles⁸ that this very fact evidences how much, in the realm of grammatical relationships, syntax itself has become dependent on semantics.⁹

As regards the lexicon, it seems quite evident that, as one deals with features related to semantics and their impact on syntax, one is dealing in fact with the issue of categorizing. Assuming that grammatical categories and processes cannot be separated from semantic roles, it is relevant for syntax to look at grammatical categories, their structure and the kind of information they convey. This means redefining syntax and shaking the very foundations of lexicon, where categories have always been fundamental.

Which ever stance one adopts when looking at the lexicon, at the current stage of linguistics, one will find that there is an association between grammatical aspects and categorizing. On the other hand, categorizing — vital for a lexicon, inseparable from grammatical relationships — is a crucial point to the extent that, again at the current state of linguistics, categories are no longer taken as preset; they are no longer determinable on the basis of formal criteria, nor do they have equivalent properties. Categories here are linked to conceptual models, which means they are given linguistic focus, tied to a certain way of encoding experience. In view of this, it is not difficult to understand why, in developing a dictionary which reflects lexical aspects of a language, one can no longer rely only on linguists and their data, and/or the overlapping data from occasional collaborators.

The four aspects of lexicon discussed here all require dialogue between all would-be participants in the dictionaryization process; only then, can a dictionary be compiled.

EFFECTIVE PARTICIPATION BY THE INDIANS

In the specific case of the Tikuna Indians, native speakers were able to participate effectively because they were encouraged to reflect on their language¹⁰ at several levels:

- a) the text production and collection level;
- b) the classification of texts not only according to source but also the social role of the text's author, classification that leads speakers to operate with their own concepts of verbal genre and cultural domain;
- c) recognition of words and phrases in oral texts;
- d) extraction of meaning — from words and sense relations;
- e) working with semantic wholes related to structures larger than words: here semantic and syntactic roles can be articulated;
- f) categorizing pitch: in a tonal language like Tikuna, words can be correlated on the basis of their tonal patterns;
- g) text transcription, in which one can check whether the categories established in the oral text undergo any changes when mediated by writing;
- h) structuring the dictionary and structuring entries.

The participation of native speakers is thus combined with the linguist's activity. Since the activities of all concerned are interconnected, we have referred to these as an action that *speaks with*. Its scope will vary according to the angle(s) from which words and their meanings are approached.

Having defined an action that *speaks with* and described the effective participation of native speakers, we are led to discuss the relationship between dictionaryization and translation.

BRIEF REMARKS

The participation of Tikuna speakers — unlike what is usually the norm regarding Indian languages — will result in a dictionary of much broader scope than usual. In view of its underlying concepts, the dictionary will cover: definitions, categories, linguistic usage variations, grammatical features, homonymy, polysemy and paronymy, synonymy, pronunciation.

Since it is conceived as a source of reflexion, an inventory of established signs, and a means of visualizing lexical regularities and establishing correlations, the dictionary would have subdivisions covering three types of correlations: 1) correlation between meanings; 2) correlation between forms; 3) correlation between sound elements.

The first type includes synonymy and metaphor, for example. The second involves forms which originate in certain cultural realms. The third deals with certain tonal patterns.

The proposed dictionary is not intended to verify words. As is evident, it does not fall into the pattern of traditional lexicography ratified by an Academy or similar entity. However, the breadth of the dictionaryization process envisages a linguistic approach that would produce a dictionary quite similar to a mother tongue dictionary. Such breadth, however, does not exclude the production of a bilingual dictionary within the very process of dictionaryization. It has been shown that good bilingual dictionaries are supported by mother tongue dictionaries.¹¹ Therefore, for an Indian language, such as Tikuna, which has no mother tongue dictionary, the process of dictionaryization may certainly lead to a good bilingual dictionary. A bilingual dictionary would contain translations of Indian words and phrases into another language. Obviously, there would be a permanent risk of endorsing words and phrases of the other language, but this risk is minimal in our proposal. It has been said that in order to translate¹² one must experiment with "fluidity" and learn to attribute lexical meaning in a language whose logical and semantic framework is quite different from that of the language into which the translation is being done. We might add that in order to translate terms from an Indian language one must go beyond comparisons based on the linguist's analysis. In our proposal, due to the many levels of effective participation by the Indians themselves, the dictionaryization process will be based on the perspective of the Indian group whose language is at stake. As a result, there is reduced risk of nonvalidation of the terms translated from an Indian language into another language where those terms might not reflect identical realities. Consequently, this will increase the chances of the Indian group feeling that the material product belongs to them since it embodies their knowledge.

Notes

1. See Wiesemann (1981), who says in the preface to her dictionary: "This dictionary has been conceived with the main purpose of providing the Kaingáng Indians with more help in their endeavors to learn the Portuguese language..."
See also Emiri (1987: 10): "...I wanted [...] to write the result of processing the data collected about the Yānomamè thematic universe to [...] give back to the Yānomamè their knowledge that I have merely set in a systematic format, in order to help them become aware of, assert and protect their own ethnic identity..."

2. We are thinking here of Indian teachers.
3. This includes, for instance, polysemy, synonymy, antonymy, hyponymy and paraphrases.
4. A simplifying criterion, for example, would be working with the Bloomfieldian concept of word.
5. There should be no confusing the statement that a lexicon has words with different backgrounds and the neogrammatical notion of change of meaning, in which the process of change is seen as an adventure of forms individually considered, with no reference to rival forms.
6. Cf. Bloomfield (1993: 426-427).
7. The autonomous syntax line is followed, for example, by Chomsky (1981) and Marantz (1984).
8. An example of such dependence is found in Marantz (1984), where the notions of logic object and logic subject, which are dependent on the definition of semantic roles, supply the essential basis for subcategorizing verbs within a certain lexical category.
9. The kind of semantics that has led syntax to incorporate semantic roles is that which has moved from a concern with the meaning of a sentence as being represented in a phrasal tree into a concern with meaning as relative to conceptual scenes (including here the meaning of the sentence). On this subject, see works by Fillmore listed under References.
10. In the context of Brazilian Indian groups, the Tikuna have been remarkable for having a strong participation in preparing the teaching materials used within their own group.
11. We reached this conclusion after analyzing and comparing a good number of native language dictionaries and bilingual dictionaries involving the so-called Western languages, in addition to dictionaries on Brazilian Indian languages.
12. See Camargo (1991: 65).

REFERENCES

- BASÍLIO, M. (1987): *Teoria lexical*, Rio de Janeiro, Ática.
- BLOOMFIELD, L. (1933): *Language*, London, George Allen & Unwin Ltd.
- CAMARGO, E. (1991): *Phonologie, morphologie et syntaxe. Étude descriptive de la langue Caxinauá (Pano)*, Thèse présentée en vue du grade de Doctorat de Nouveau Régime, Paris, Université Paris IV — Sorbonne.
- CHOMSKY, N. (1981): *Lectures on Government and Binding*, Dordrecht, Foris Publications.
- EMIRI, L. (1987): *Dicionário Yãnomamê-português (dialecto Wakatautheri)*, Boa Vista, Edições CPI/RR.
- FILLMORE, C. J. (1968): "The Case for Case", Bach, E. and R. Harms (Eds.), *Universals in Linguistic Theory*, New York, Holt, Rinehart and Winston.
- FILLMORE, C. J. (1976): "Topics in Lexical Semantics", Cole, R. (Ed.), *Current Issues in Linguistic Theory*, Bloomington, Indiana University Press.
- FILLMORE, C. J. (1977): "The Case for Case Reopened", Cole, P. and J. Sadock (Eds.), *Syntax and Semantics, vol. 8: Grammatical Relations*, New York, Academic Press.
- LAKOFF, G. (1982): *Categories and Cognitive Models*, Cognitive Science Program, University of California at Berkeley.
- LANDIN, D. (1983): *Dicionário e léxico Karitiãna-português*, Brasília, S.I.L.
- MARANTZ, A. (1984): *On the Nature of Grammatical Relations*, Cambridge, The MIT Press.
- WIESEMANN, U. (1981): *Dicionário Kaingáng-português, português-Kaingáng*, Brasília, S.I.L.