Foregrounding the Development of an Online Dictionary for Intermediate-level Learners of Brazilian Portuguese as an Additional Language: Initial Contributions

Tanara Zingano Kuhn

Abstract

The present PhD project intends to collaborate with the designing of a monolingual online dictionary for intermediate-level learners of Brazilian Portuguese as an additional language. Considering that the development of such a reference work involves the investigation of a series of theoretical-methodological aspects, this research will be narrowed down to one specific issue: the use of simplified Portuguese language patterns in the writing of the definitions. Therefore, the steps to be taken entail a thorough bibliographical review on lexicographical definitions for monolingual learners' dictionaries and the use of defining vocabulary for their writing; Brazilian Portuguese corpus research in order to compile a defining vocabulary list (DVL); and tests with learners to verify which kind of definitions – those which were written with or without the use of DVL – is better for the user. Since pedagogical (meta)lexicography regarding Brazilian Portuguese as an Additional Language (BPAL) is to a fairly large degree still incipient, especially when compared to what has been done in the area of English as a Foreign Language (EFL), this project is expected to give substantial contribution to new knowledge.

1. Background

In April 2010 the project Ensino de português como língua estrangeira: bases para um dicionário[1] on-line para suporte de atividades em EAD [*Teaching Portuguese as a foreign language: foundations for an online dictionary as an auxiliary tool for activities in distance learning*] officially began after approval by the Distance Education Office at the Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil. The team that develops such a project consists of Professor Maria José Bocorny Finatto (creator and coordinator of the project); Professor Margarete Schlatter (founder and coordinator of the Centre of Portuguese for Foreigners (PPE) at UFRGS); Tanara Zingano Kuhn, PhD student in the Department of Linguistics at Leiden University and former full-time lecturer of Portuguese at Hankuk University of Foreign Studies (Seoul, South Korea); Aline Evers, M.A. student at UFRGS and teacher of Portuguese as an additional language; and Lucas L. Iochpe R. Guerra, undergraduate student of Computational Sciences who is financed through a scholarship quota awarded by UFRGS.

Dictionaries of Portuguese language first came out over 200 years ago. However, the process of solidification in a written form of the linguistic norm and vocabulary of Brazil began much later, and this can be justified due to the fact that the dictionaries produced in the XIX century referred to European Portuguese only (Biderman 2003).

The first dictionary to represent the Brazilian Portuguese (BP) variety, the *Pequeno Dicionário brasileiro da língua portuguesa*, was published in 1938. Since then, Brazil has produced a sheer number of general dictionaries and what we call 'dictionaries for school learners', which are geared towards Brazilian students at elementary and secondary schools. Besides those, there is also a wide variety of bilingual dictionaries for Brazilian learners of foreign languages.

Despite the apparent success that Brazil has had in validating its language variety in the Lusophone world, there is still a gap to be bridged as to the dissemination of our language

to foreign learners: a monolingual dictionary of Brazilian Portuguese for foreign learners has never been published.[2]

The need to produce this material is justified not only in views to satisfy a market demand but, first and foremost, it is aimed at being a qualified pedagogical tool. We agree with Yamada (2009:148), 'I consider learner's dictionaries as resources with wide-ranging applications for language learning and teaching rather than as mere reference tool'. Hence, we expect the increasing number of people who are learning Brazilian Portuguese all over the world and teachers alike to benefit from the material we are developing. In addition, the study involved in the production of this dictionary will certainly contribute to the area of lexicographical pedagogy, a still quite incipient field in Brazil, notably concerning dictionaries for foreign learners of BP.

In this regard, the present research project intends to collaborate with the development of the abovementioned dictionary. Among a series of theoretical-methodological measures that dictionary making involves, I have chosen to further study the lexicographic definition since a well-structured and solid foregrounded defining strategy seems to pave the way for the production of a highly-qualified material; besides, such a study has never been done for Brazilian Portuguese as an additional language.

2. Purpose of the study

Considering that 'definition lies at the heart of lexicographic practice, being the central part of the microstructure of a dictionary entry' (Fabiszwesky-Jaworsky, Grochocka 2009:90), lexicographers tend to invest a great deal of time trying to elaborate them according to the precepts of Monolingual Learner's Dictionaries (MLDs). According to Hornby (apud Bogaards 2009:12), the three main characteristics of this kind of dictionaries are: a selected vocabulary, simple definitions, explicit information about use. However, doing so is not an easy task. Peter Sokolowski, when talking about how Merriam-Webster's Advanced Learner's Dictionary was created from scratch, shared with the readers how his team of lexicographers felt during that process: '...creating the simplest possible definitions for the English language learner presented an enormous challenge over the course of this project' (2009: 48). The elaboration of lexicographic definitions for the online dictionary for learners of Brazilian Portuguese as an additional language shall follow suit. However, we might benefit from the fact that we are starting this project from scratch exactly 70 years after the first monolingual English dictionary for learners was ever made (*Idiomatic and Syntactic English Dictionary* (ISED) was published in Tokyo in 1942, see Bogaards 2009): we can learn from their experiments, adopt the good results and adjust the not so good ones.

2.1. *General purpose*

The research question that structures this project is: *how to write definitions in a plain, simple and unambiguous way*[3]?

We expect to find some hints on how to do that through a thorough bibliographical review of previous work in the area. This entails not only a reassessment of the 'Big Five' (that is, the five worldwide known MLDs of English; namely, *Cambridge Advanced Learner`s Dictionary*, *Collins Cobuild English Dictionary for Advanced Learners*, *Longman Dictionary of Contemporary English*, *Macmillan English Dictionary for Advanced Learners*, and *Oxford Advanced Learner`s Dictionary*) and MLDs in other languages, but also a

reappraisal of the theoretical and experimental studies that researchers have been doing since the publishing of ISED in 1942.

These are three fundamental questions that need to be answered, comprising both the content of the definition (1) and its form (2, 3):

1) Bearing in mind our target user`s profile and their needs (encoding/decoding purposes), what kinds of and how much information should we include in the definitions? How to decide on the number of lexical units to be instantiated? How should we deal with extralinguistic messages (pragmatics, sensitivities, connotation) (see Atkinson and Rundell 2008)?

2) What kind of defining style should we adopt: classic (analytic, Aristotelian) definition, single clause "-when" definition, contextual (full-sentence, COBUILD-style) definition or synonym definition (see Svénsen 2009)? Should we opt for just one of them or can we adopt various types according to the part of the category of the headword in question?

3) What kind of words can we use in the definition? Should we base on lexicographers instinct or should we have the help of a defining vocabulary[4] to determine which words can and cannot be used?

We can try to find the answer for questions 1 and 2 through the review of previous work. After reassessing this material, we can get to a final proposal, which should be included in the Style Guide of our dictionary to be followed by future lexicographers involved in the project. Question number three, however, imposes further measures, because there is not a pre-established defining vocabulary list for Brazilian Portuguese.

2.2 *Specific purposes*

The use of defining vocabulary (DV) for the writing of definitions in Monolingual Learner's Dictionaries (MLDs) has been the focus of investigation and debate of many researchers. Some authors claim that the application of such a method brings forth significant disadvantages, while others affirm that the overall positive results obtained through the employment of this strategy surpass the drawbacks.

Rundell (1998) alerts that the use of DV might, for example, hinder the process of writing for the lexicographers, due to 'arbitrary constraints on lexicographers' freedom to define' (319). Consequently, unnatural forms of expression in definitions end up being formulated, as Svénsen (2009) points out: 'Some definitions have been criticized for using a language that few language teachers would wish their students to imitate'(248). As a brief sum-up, Yamada (2009: 151), when presenting an analysis of EFL dictionary evolution, refers to the work of Kawamura (2009: 87-89) to list six difficulties that concern defining vocabulary:

1. Inclusion of lexical items beyond the expected proficiency of EFL dictionary users

2. Lengthy definitions

3. Unnatural definitions

4. Senses to be used are not controlled

5. Actual size greater than advertised

6. Actual use of defining vocabulary is unclear

On the other hand, Herbst (1986) analyses the defining vocabulary policy used in *Longman Dictionary of Contemporary English* (LDOCE), and, despite its inadequacies and inconsistencies (he actually mentions those drawbacks above), concludes that this is a valuable instrument. The author sustains his argument, firstly, through the presentation of tests results, which prove that:

> The LDOCE approach has succeeded in providing definitions that, with some reservations, are sufficiently accurate, but which are definitely written in a much simpler language than those in ALD and will thus be considerably easier to interpret to the foreign learner (Herbst 1986:112).

Secondly, he points out that the use of a controlled vocabulary, in fact, helps lexicographers do their work, because such a list contains the easiest words for a foreigner to understand; hence, words outside this list are considered to be difficult and should be avoided. The obligation to fit to this pattern of writing makes lexicographers be more aware of their writing, or, as Herbst puts it, 'inforce lexicographical discipline'(113). The result is, according to the author, a fundamental contribution to the final goal of simple comprehensibility.

From this brief review one can conclude that the use of DV with reference to the English language is a controversial topic. As to Brazilian Portuguese, we can only suppose this premise should follow. In order to analyse the pros and cons of using a controlled vocabulary in the writing of the definitions in our dictionary, we should run tests with learners. Ideally, the results would help determine the adoption or not of this method for our final work.

However, in order to run those tests with learners and verify which kind of definition is better, the one that uses defining vocabulary or the one that does not, we need to have access to a defining vocabulary list, which does not exist for Brazilian Portuguese[5]. We must, first of all, elaborate such a list. The following questions must be taken into consideration:

1. Considering that defining vocabulary lists are obtained through corpus analysis, what corpus should we use in order to elaborate a defining vocabulary of Brazilian Portuguese? Are the existing BP corpora balanced, that is, do they 'reflect the diversity of the target language, by including texts which collectively cover the full repertoire of ways in which people use the language' (Atkins and Rundell 2008: 66)?

2. Does the fact that the target-user has an Intermediate level of proficiency in Portuguese matter in the choice of the corpus? If yes, how can we select the corpus to be used? If no, how can we guarantee the words selected for the defining vocabulary will be known by the target-users?

3. How can we determine the size of the defining vocabulary? Should the list be composed of lemmas or lexical units? Should we add a list of profitable affixes (like in *Longman Dictionary of Contemporary English*) or should we present the words in the full form?

3. Methodology

The methodology of this research involves four stages, being the first one of a theoretical nature and the other three of a practical one:

1) Bibliographical review of previous work in the area of lexicographical pedagogy, especially to what concerns MLDs. I will reassess the Big Five and other dictionaries

of languages for foreigner learners, alongside with a critical reading of texts from authors whose work is related to this subject. Furthermore, as to learner`s needs, I will study the profile of our target-user in order to tailor the writing of the definitions to their proficiency level and need;

2) Compilation of a defining vocabulary list through corpus investigation. This stage involves corpus analysis through computational programs like Word Sketch Engine and Palavras;

3) Writing of two sets of definitions: one through the use of the defining vocabulary list, the other one without. Here we need to carefully choose which headwords (abstract or concrete; one or more parts of the speech) will be defined so that our tests will not be hindered by unexpected variables. Also, we will need to follow the defining style we previously selected as the most appropriate and bear in mind our users` need;

4) Testing with users. We intend to run the tests with Intermediate-level learners of BPAL that are studying at PPE-UFRGS[6] . Since they have different nationalities, we might separate the groups according to their mother-tongue and knowledge of an additional language other than Portuguese, as an effort to avoid that these variables interfere with the results. This stage of the research must be carefully planned and rely on statistics knowledge.

4. Final remarks

The objective of this paper was to present my PhD research as a study that seeks to collaborate with the development of an online dictionary for intermediate-level learners of Brazilian Portuguese as an additional language through the investigation of lexicographic definitions. I presented the background of this project; its general and specific purposes; and the methodology to be followed. Since this research is in its very initial phase, results could not be obtained yet. I hope the ideas here exposed motivate those who are interested in Brazilian Portuguese language and pedagogical lexicography to help foreground this enterprise.

Notes

[1] http://www6.ufrgs.br/letras/dicionarioportuguesle/

[2] There is one book, the *Frequency Dictionary of Portuguese – Core vocabulary for learners* (by Mark Davies and Ana Preto-Bay), which is explicitly aimed at learners; however, this material presents both varieties of Portuguese (European and Brazilian) and next to the headwords in Portuguese they provide the equivalent in English followed by examples in Portuguese. Therefore, it is not a monolingual dictionary and it is not exclusive of Brazilian Portuguese.

[3] Henry Sweet, already in 1899 (apud Herbst 1986:101), affirmed that 'The first business of a dictionary is to give the meanings of words in plain, simple, unambiguous language'. This recommendation actually referred to general dictionaries; however, it seems this principle fits perfectly well to learner's dictionaries.

[4] 'A systematically selected range of words to be used for describing the content of a larger number of words is usually called a (CONTROLLLED) DEFINING VOCABULARY.' (Svénsen 2009: 246)

[5] For an initial investigation on the quest of producing a defining vocabulary list, see Kuhn, T., Finatto, M.J. (2011) *On the proposal of an on-line Brazilian Portuguese dictionary for speakers of Asian languages: An ongoing experiment*. Proceedings of Asialex 2011. P.284-294

[6] PPE stands for *Portuguese for Foreigners* and is a prestigious centre of BPAL teaching in Brazil. It is part of UFRGS and receives, every year, hundreds of new students.

References

Atkins, B.T.S. and M. Rundell 2008. *The Oxford Guide to Practical Lexicography*. New York: Oxford University Press Inc.

Biderman, M.T.C. 2003. 'Dicionários do Português: da tradição à contemporaneidade. [Dictionaries of Portuguese: from the tradition to present day].' *Alfa* 47.1: 53–69.

Bogaards, P. 1998. 'Scanning Long Entries in Learner`s Dictionaries.' *Actes EURALEX`98 Proceedings* II, 555–563.

Bogaards, P. 2009. 'The evolution of learner`s dictionaries and Merriam-Webster`s Advanced Learner`s English Dictionary.' *English Learner`s Dictionaries at the DSNA 2009*, 7–27.

Fabiszewski-Jaworski, M. and M. Grochocka 2009. 'Folk defining strategies vs. comprehension of dictionaries definitions: an empirical study.' *English Learner`s Dictionaries at the DSNA 2009*, 89–105.

Herbst, T. 1986. 'Defining With a Controlled Defining Vocabulary in Foreign Learners`Dictionaries.' *LEXICOGRAPHICA*, 101–119.

Rundell, M. 1998. 'Recent Trends in Pedagogical Lexicography.' *International Journal of Lexicography* 11.4: 315–342.

Rundell, M. 2006. 'More than One Way to Skin a Cat: Why full sentence definitions have not been universally adopted.' In E. Corino, C. Marello and C. Onesti (eds.), *Atti del XII Congresso Internazionale di Lessicografia : Torino, 6-9 settembre 2006*. Alessandria: Edizioni dell'Orso, 323–338.

Sokolowsky, P. 2009. 'From the ground up: Making a new learner's dictionary from scratch.' *English Learner`s Dictionaries at the DSNA 2009*, 45–54.

Summers, D. 1993. 'Longman/Lancaster English Language Corpus – Criteria and Design.' *International Journal of Lexicography* 6.3: 181–208.

Svénsen, B. 2009. *A Handbook of Lexicography. The Theory and Practice of Dictionary-Making.* Cambridge: Cambridge University Pres.

Yamada, S. 2009. 'EFL dictionary evolution: Innovations and drawbacks.' *English Learner`s Dictionaries at the DSNA 2009*, 147–168.