

Do We Need Another Theory of Lexicography?

HENNING BERGENHOLTZ

1. Introduction

The title of this contribution was inspired by Tono (2010, 2), who asks the following questions about the future of lexicography:

- (i) Do we really need a “theory”?
- (ii) Why bother to construct another theory of lexicography?
- (iii) Can lexicography be really an independent scientific discipline?

Initially, the answers Tono (2010) gives to questions (1) and (2) are somewhat sceptical, but in the course of the contribution the answers turn distinctly positive. As far as I understand the contribution, it does not offer a straight answer to the third question. All three of these questions are highly relevant to the future development of the theory and — if one assumes that theory has a material influence on practice, as I do — of the practice of lexicography.

Not to beat about the bush: my answer to the first two questions is a resounding yes. The third question requires a somewhat longer, but nevertheless still unequivocal answer: Lexicography requires the collaboration of different experts, depending on the type and content of the dictionary (see Bergeholtz and Gouws (To appear)); for example, the design and compilation of an English-Spanish gene technology lexicon requires one or more molecular biologists, one or more technical linguists for English and Spanish with a sound knowledge of molecular biology, and one or more lexicographers. For a dictionary of classical music, one or more experts on music as well as one or more lexicographers are needed. In all lexicographic projects a lexicographer is an

essential and leading collaborator. At its core, lexicography is a genuine subset of information theory, but it is a special one that has specialised in the processing and presentation of data (see Bergholtz and Bothma (2011) in this regard). We therefore do not subscribe to the widespread classification of lexicography as a subdiscipline of linguistics.

2. Lexicography as applied linguistics

There is no doubt that the present situation must be seen as a result of its context and its history. Before around 1960, lexicography did not exist at universities as a theoretical discipline. Since the 15th century any number of theoretical considerations have appeared, especially in the prefaces to dictionaries, lexicons and encyclopaedias. But it was not an academic discipline with its own chairs, journals and international associations. In the years thereafter, two parts of lexicography developed into scientific disciplines:

1. Specialized lexicography (it calls itself terminology or terminography); the scientists involved were and are mostly linguists (but not only linguists, subject specialists were and still are involved).
2. For language lexicography (interpreted as lexicography for communicative dictionaries and their theory) a discipline developed in the 70s and 80s that called itself lexicography and initially defined itself as applied lexicography.
3. No comparable academic discipline developed for encyclopaedias and lexicons. Those who have called themselves lexicographers did make limited use of such reference works, but they did it in a step-motherly fashion rather than systematically.

This situation proved to be untenable, especially with reference to Internet dictionaries. Besides the small group of terminologists, and in the absence of all encyclopaedists, linguistics laid claim to the dictionary arena and its theory. However, this has not been successful, as I would like to explain with reference to two Danish examples. I must point out that these Danish examples are typical of the semi-theoretical status of lexicography and the associated disadvantages in the scientific

world. In journals such as the *International Journal of Lexicography* or *Lexicon* linguistic problem areas are often discussed, but without real linguistic originality. When such journals publish novel contributions, they offer lexicographic novelty — which is hardly likely to impress a linguist when he assesses the value of the new thinking to linguistics.

Case 1: In the process of a new system for allocating research funds, journals were classified in Denmark according to status, i.e. the assumed quality of research in the different journals. Whoever publishes in the best journals sees to it that the university, faculty or institute gets a larger slice of the total amount to be distributed. To this end a subject typology with an allocation to subdisciplines was drawn up. The classification of the journals into a higher or lower group was done by subject representatives appointed for this task. As lexicography was assigned to linguistics, and there was no lexicographer on the committee, the lexicographic journals did not make it to the highest level. From a linguistic point of view, this is understandable: lexicographic contributions do not propose new and therefore important linguistic theories. This allocation has enormous disadvantages for lexicography. We cannot publish high-level articles if we do not also publish contributions to another scientific discipline.

Case 2: The Center for Lexicography asked the dean of its faculty to sign a co-operation agreement with a foreign university that provided for collaboration between two research groups. The object of the intended research was to design and to actually produce a bilingual dictionary of fixed expressions. As explained in the application, new concepts for lexicographic databases and the relation between databases and search routes, search steps and search times on the one hand and electronic dictionaries on the other would be key elements of the joint research and would subsequently be published in relevant international journals. The University's response to this request was negative. It was said that such cooperation would not be appropriate at university level; it was very concrete, since its object would be a physical dictionary. The university management apparently viewed lexicography as a part of linguistics and stated that no linguistic theories and no novel linguistic

empirical research were foreseen. On the contrary, the objective was a concrete practical tool, namely one or several dictionaries. It was concluded that this was a practical and not a theoretical project and therefore not a university project.

3. Do we need lexicographic theories at all?

The answer to this question depends on who “we” are. If equally good lexicographic products can be developed without underlying theories as with them, then one probably doesn’t need them. Then lexicography would indeed not be a university discipline. Lexicography would be an “art and craft”, but not a science. Since many lexicographers often (or also) work at universities, there are at least three possible arguments that “save” lexicography as a university subject. The first argument is the customary one, represented especially in British lexicography:

1. Lexicography does not have its own theory, and it does not need one either, as it can resort to existing linguistic theories and needs no further theories, according to Atkins and Rundell (2008, 4) or Bejoint (2010, 161): “There are theories of language, there may be theories of lexicology, but there is no theory of lexicography.”
2. Lexicography has its own theory that consists in describing structures of existing dictionaries; this view is held by Wiegand (1998) and other, especially German lexicographers. Actually, Wiegand’s theory is a linguistically oriented structure description, according to the critique of Bergenholtz and Tarp (2003).
3. The modern functionalist theory. This is the theory that is the object of the contribution by Tono (2010). It is his observations I wish to comment on below.

I would like to stress in advance that although I will question some of the views expressed by Tono, on the whole I regard his contribution as very important to the future of practical as well as theoretical lexicography. Generally speaking, I view his presentation as fair and correct. It is the points about which we do not quite agree which are interesting and call for further study. And these are the points I want to examine

more closely below.

4. Users know nothing about lexicographical theories

Somehow it seems unrealistic to assume that someone using a saw will wonder whether empirical research has been carried out on the uses of saws, or whether scientific research has been conducted about the choice of material for or design of a saw. The user looks at the price and checks whether the description of the intended uses of the saw correspond to these needs — or he asks the salesperson on the floor. The same applies to the tool called a dictionary. In this regard Tono (2010, 6) is right when he says that “people usually never call for a theory in dictionary making after all”. This is of little or no interest to the average dictionary user. It may boost sales if a well-known professor or a famous actor praises the dictionary. The user resorts to a dictionary because he has a problem which he thinks the dictionary can help him solve. If the dictionary supplies an answer, he trusts it; if not, he gets annoyed. Just like the user of a saw gets annoyed if his saw cuts too coarsely or cannot cut the tree or the board at all because it is not strong enough. Theories are things one can benefit from in order to manufacture certain tools: there is nothing more practical than a good theory. However, if — like Tono (2010, 6) — one says of the public opinion that theories have “associative links with the hard sciences”, then that is of no interest to any science that has any business being practised at a university. Here one needs, and has, theories in both hard and soft sciences, otherwise they would not be scientific disciplines.

5. Dictionaries are products

Dictionaries are products (Tono 2010, 13). Quite so; but more precisely: products that can serve as tools. If they sell well, such products are available in constantly new, more or less modified versions. In the case of internet dictionaries, the changes and additions can be inserted every day or even every minute. This should be done, and is in fact being done — also in the various internet dictionaries we have produced in Denmark. But this is not done, like Tono (2010, 15f) claims, on the

basis of user surveys that have been conducted, but through constant editorial deliberations and feedback from individual dictionary users, who send e-mails to the editors. Such feedback, which can also be called user surveys as far as I am concerned, is never of a theoretical nature, but always very concrete. For example:

1. The following word or the following meaning is missing. Could you add this word or meaning to the dictionary?
2. You have the following entry . . . In my view this is incorrect, because . . . I suggest you change the description in this case.

Such feedback is combined with log file analyses of various kinds (see Bergenholtz 2006, Bergenholtz and Johnsen 2007, 2011). Among other things, one can see what words not listed in the lemma stock are searched for and insert them into the dictionary. In such cases, feedback and log file analyses have concrete effects on concrete dictionary entries. In other cases, they can also bring about a new dictionary concept. For example, with *Vrang/Bergenholtz/Almind (2002–2006)* we had a dictionary of idioms with no less than 9000 entries. Nevertheless, we kept receiving mails from users requesting us to include in the dictionary turns of phrase which by our classification were not idioms, but proverbs, for instance, or phrase verbs. Moreover, log file analyses revealed that around 27% of all dictionary searches failed because the search object was not an idiom. Because of these findings, the old dictionary of idioms was withdrawn, and a new dictionary design was developed and produced that contained all forms of fixed phrases (Bergenholtz/Bjærge/Almind 2010a, 2010b, 2010c, 2010d).

This “dictionary” therefore comprises not one dictionary, but four, all of which grew out of the same database of fixed expressions. There are four monofunctional dictionaries — three communicative dictionaries and one cognitive dictionary (see Bergenholtz 2011):

1. Reception dictionary (Bergenholtz/Bjærge/Almind 2010a).
2. Production dictionary where the search string is an expression (all or part of a fixed expression) (Bergenholtz/Bjærge/Almind 2010b).

3. Reception dictionary (Bergenholtz/Bjærg/Almind 2010a).
4. Knowledge of fixed expressions (Bergenholtz/Bjærg/Almind 2010a).

The idea is the simple one that tools should always be fit for purpose if they are to give optimal assistance. For reception problems, one needs only an explanation of the meaning and nothing more. Indirectly, this was also based on mails from dictionary users in combination with own experiences of other dictionaries (Bergenholtz 2011). Many dictionary entries had become so long that the users were no longer able to find certain existing data. The users had terminated the search because of information overload.

The principle is therefore not, as Tono (2010, 13) put it: “It would be very interesting if all the different dictionary contents can be searched via a user based function-based menu.” As I see it, the relevant question is different: How can one construct a dictionary database that contains all the data types and which the users can search specifically so that they can easily, quickly and reliably find answers to different concrete types of questions? For the production dictionary, the previously defined search consists in a set search sequence in certain database fields together with a presentation of data types in a set sequence (for a more detailed explanation see Bergenholtz 2011):

Fields searched+order of search	Field	Order in dictionary entry
1	1. Core field	
	2. Meaning	2
	3. Further meaning item	3
	4. Grammar	4
	5. Remark(s)	
	6. Internet link	
2	7. Fixed expression(s)	1
	8. Style	
	9. Classification of the fixed expression	

3	10. Collocation(s)	5
4	11. Example(s)	6
	12. Synonym(s)	7
	13. Antonym	8
	14. Association(s)	

This is a search and a data presentation that has been specified for the text production function by the lexicographer. And from the database one can extract only four dictionaries (at present). “How far can function go?” Tono (2010, 21) asks. My answer is: As far as one would like to; as far as any individual user would like to. It is possible — and this is what we are planning for this database — to allow every individual user to specify such settings and even to self-define a series of search combinations and display options. This should then be possible not only by means of simple search strings, but also by using Boolean operators. In principle, thousands of dictionaries can then be defined which are based on this relatively simple database. It does not have to be done only in accordance with the functionality dictated by the lexicographer. I assume that such advanced searches would be done by only a relatively small percentage of dictionary users, but this is not a key issue. Lexicographers can offer users such functionality without great effort.

6. Lexicography as part of information science

Classifying lexicography as a part of information science could spell the beginning of the end of the underestimation of lexicography. With a few slight changes or some fine-tuning, the figure presented by Tono (2010) could be a good starting point. One refinement could be that “user research” should not be interpreted as user surveys, but that user reactions should also be considered; and it would also be advantageous to take certain experiments conducted with only a small number of respondents into account. Another refinement could be the replacement of “confirm/reject hypothesis” with “reject parts of the theory or correct

a certain dictionary conception". I do not think that it is possible to obtain verification of modern functional theory or even of parts of it. But it is certainly possible to identify uncertainties and outright errors of judgment, such as the lemma lacunas described above in a concrete dictionary or the inadequate design of the dictionary of idioms referred to above, through user feedback and log file analyses, which led to concrete changes, supplements to individual entries or a change of the overall concept. Such changes in concrete dictionaries can then lead to basic supplements or changes of a theoretical nature, as happened in the case of the plea for monofunctional dictionaries.

NOTE

I want to thank Theo Bothma (University of Pretoria) and Mei Xue (University of Aarhus) for interesting discussions on Yukio Tono's contribution in *Lexicon* 40 (2010).

REFERENCES

- Atkins, Sue and Michael Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Béjoint, Henri. 2010. *The Lexicography of English*. Oxford: Oxford University Press.
- Bergenholtz, Henning. 2006. "Idiomwörterbücher und ihre Benutzer". In *Wörter — Verbindungen. Festschrift für Jarmo Korhonen*, hrsg. von Ulrich Breuer und Irma Hyvärinen. Frankfurt a.M., Bern, New York and Paris: Peter Lang, 19–30.
- Bergenholtz, Henning. 2011. "Access to and presentation of needs-adapted data in monofunctional internet dictionaries". In P. A. Fuertes-Olivera / H. Bergenholtz (eds.): *e-Lxicography: The Internet, Digital Initiatives and Lexicography*. London & New York: Continuum. (in print)
- Bergenholtz/Bjærge/Almind. 2010a. = Henning Bergenholtz under medvirken af Esben Bjærge: *Betydning af faste vendinger*. Database og design: Richard Almind. Odense: Ordbogen.com. (www.ordbogen.com)
- Bergenholtz/Bjærge/Almind. 2010b. = Henning Bergenholtz under medvirken af Esben Bjærge: *Brug af faste vendinger*. Database og design: Richard Almind. Odense: Ordbogen.com. (www.ordbogen.com)
- Bergenholtz/Bjærge/Almind. 2010c. = Henning Bergenholtz under medvirken af Esben Bjærge: *Faste vendinger med en bestemt betydning*. Database og design: Richard Almind. Odense: Ordbogen.com. (www.ordbogen.com)
- Bergenholtz/Bjærge/Almind. (2010d). = Henning Bergenholtz under medvirken af Esben Bjærge: *Betydning af faste vendinger*. Database og design: Richard Almind. Odense: Ordbogen.com. (www.ordbogen.com)
- Bergenholtz, Henning and Theo J. D. Bothma. 2011. "Needs Adapted Data Presentation in e-Information Tools". *Lexikos* 21 (sent to the journal for peer review November

- 2010)
- Bergenholtz, Henning and Rufus Gouws. To appear. "Who is a lexicographer?" *Lexicon*.
- Bergenholtz, Henning and Mia Johnsen. 2007. "Log files can and should be prepared for a functionalistic approach". *Lexikos* 17: 1–20.
- Bergenholtz, Henning and Mia Johnsen. 2011. "User research in the field of electronic dictionaries: Methods, first results". In *Dictionaries. An international encyclopedia of lexicography. Supplementary volume: Recent developments with special focus on computational lexicography*. Ed. by Rufus H. Gouws, Ulrich Heid, Wolfgang Schweickard, and Herbert Ernst Wiegand. Berlin: Mouton de Gruyter: Berlin. (in print)
- Bergenholtz, Henning and Sven Tarp. 2003. "Two opposing theories: On H. E. Wiegand's recent discovery of lexicographic functions". *Hermes, Journal of Linguistics* 31: 171–196.
- Tono, Yukio. 2010. "A critical review of the theory of lexicographical functions". *Lexicon* 40: 1–16.
- Vrang/Bergenholtz/Almind. 2002–2006. = Vibeke Vrang med bidrag af Henning Bergenholtz og Lena Lund: *Den Danske Idiomordbog*. Database og design: Richard Almind. (www.idiomordbogen.dk)
- Wiegand, Herbert Ernst. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. 1. Teilband. Berlin and New York: de Gruyter.

Henning Bergenholtz, *Centre for Lexicography, Aarhus School of Business, University of Aarhus, Aarhus, Denmark and Department of Afrikaans and Dutch, University of Stellenbosch, Stellenbosch, South Africa* (hb@asb.dk)