The Chicago Akkadian Dictionary: Contexts and Perspectives

UDC 811.411.11'374=111(038)
DOI 10.18485/infotheca.2023.23.1.5

ABSTRACT: The purpose of this short essay is to introduce the famed Chicago Akkadian Dictionary to the readers of *Infotheca* and recommend to their attention several aspects of the dictionary which could benefit from the insight of a mind endowed with expertise in information sciences and technologies. The paper offers a short introduction to the Akkadian language and a summary of the contexts and issues related to the *Chicago Akkadian Dictionary* project.
KEYWORDS: Akkadian, dictionary, digitization.

PAPER SUBMITTED: 12 July 2023
PAPER ACCEPTED: 25 July 2023

Marko Vitas
marko_vitas@brown.edu
Brown University
Classics Department
Providence, Rhode Island, US

École Pratique des Hautes Études, Section des Sciences Religieuses
Paris, France

1 Akkadian?

Akkadian is an ancient Semitic language which originated in Mesopotamia (the region occupied, broadly speaking, by modern Iraq) and which left written traces ranging from the 3rd millennium BC down to the early 1st millennium AD. The script used to record Akkadian (and, in a preceding period, Sumerian, a linguistic isolate from the same region, as well as a number of other languages in later periods) is called cuneiform. Cuneiform is a logo-syllabic script, containing approximately 1,000 distinct signs, although not all of them were actively present during all of the periods of the script's development. Akkadian cuneiform texts are mostly preserved on clay tablets recovered from various archaeological sites across the territory of ancient Mesopotamia. The tablets, which often come down to us fragmentary and damaged, record a variety of texts: literary and mythological texts (such as the famous *Gilgameš legend* and the Babylonian poem of Creation *Enuma Eliš*) are accompanied by historical texts, such as chronicles or royal annals, religious texts and texts pertaining to the maintenance of various cults, eco-

nomical, legal (think of the *Laws of Hammurabi*) and administrative documents, letters, medical and magical texts, omens etc. When the cuneiform stopped being actively used around the beginning of the first millennium AD,[1] the knowledge of it gradually declined until it could no longer be actually read by anyone at all.[2] It was deciphered only in the 19th century thanks to the concerted effort of a number of scholars.[3]

Even this brief and simplified account of the history of Akkadian makes it easy to imagine the difficulties encountered by the readers of these ancient texts: various periods of its long history featured different and evolving vocabulary and use of language, the many genres had their own linguistic peculiarities and the fragmentary state of preservation of the texts often require the scholar to fill the missing words in the gaps.

In order to make sense of the linguistic idiom of any particular text, the scholars need to have access to the wide range of auxiliary tools, the most important of which is a comprehensive dictionary. There are today several dictionaries at the Assyriologist's disposal, among them Von Soden's *Akkadisches Handwörterbuch* in three volumes published between 1959 and 1982, and its concise English version, known as the *A Concise Dictionary of Akkadian* (known as CDA), which, in a remarkable achievement, manages to assemble a workable glossary of Akkadian into mere 450 pages. However, by far the most comprehensive and most ambitious dictionary currently available to the scholars of Akkadian is The *Chicago Akkadian Dictionary* (known as CAD).[4] This impressive dictionary, spanning 26 tomes (or 21 volumes, since a single letter sometimes required up to three tomes, as in the case of *Shin*) and almost 10,000 pages, had been in the works for almost 90 years before the final tome (*U/W*) came out in 2010. Its progress, marshaled by generations of the World's leading assyriologists and often halted and imperiled by financial concerns and acerbic interpersonal disputes, was

1. The precise timing of the script's death is debated. Geller (1997) argues that the script was still in (extremely limited) use as late as in the 3rd century AD.

2. Which does not mean that cuneiform did not continue to elicit interest among the populations which came into contact with the cuneiform inscriptions - recording various languages – which could no longer be understood in their original meaning, see (Rojas 2023).

3. Daniels (2020, 10–13) offers a good overview of the decipherment of the various languages recorded in cuneiform.

4. CAD – The Assyrian Dictionary of the Institute for the Study of Ancient Cultures of the University of Chicago.

rightly referred to, by one if its leading champions, "an adventure of great dimension."[5]

In the following sections of this essay, I will briefly sketch the history of the CAD project, as well as the dictionary's layout and content, and then turn to some of the challenges that its modern user sometimes faces in navigating such a monstrously expansive document. It is possible that some of these problems will in the future be alleviated by a more intense collaboration with experts in computer technologies, and for this reason it seemed to me that this topic could be of interest to experts in information and library sciences.

2 The Chicago Akkadian Dictionary

The most complete history of CAD comes from the memoire of Erika Reiner (entitled "An Adventure of Great Dimension"), who had been a long-standing editor of the dictionary (1973-1995). While her account of the many twists and turns in the history of CAD is – as any such account is bound to be – subjective and conditioned by her personal involvement and acquaintance with the story's main protagonists, it is a priceless source of information on the conception and development of the project, although there have been several other important accounts, at the various stages of the project, as well.[6]

The idea standing behind CAD has to be envisaged in the context of the great encyclopedic and dictionary projects which marked the late 19th and the early 20th century in Europe and North America. This was the period of great systematization in the humanities that saw the launching of the *Thesaurus Linguae Latinae* (in 1894), which is yet to be completed, the *Wörterbuch der ägyptischen Sprache* (in 1897), as well as such grand anthropological project as Frazer's *The Golden Bough* (1897-1915). In that spirit, the CAD was launched as an idea in 1921 by the esteemed assyriologist James H. Breasted, the founder of the Chicago Oriental Institute, as a comprehensive dictionary of all the many dialects of Akkadian.

The first decades of the work on the dictionary, before and shortly after the Second World War, are not very well known, even though considerable preparatory work, consisting importantly of filling out the bibliographical

5. Words of Benno Landsberger, at the annual meeting of the American Oriental Society, 1965.

6. See (Roth 2010; Stolper 1991; Gelb 1964).

cards for many words – seems to have been completed in that period. The project was, in a sense, relaunched and reinvigorated in 1954, under the directorship of Ignace Gelb, and followed through under the subsequent general editors, Leon Oppenheim (1955-1973), Erika Reiner (1973-1995) and Martha Roth (1995-2010). The 26 volumes of CAD have not been published in the alphabetic order (which would have started with A), but, taking the cue from Michigan's *Dictionary of Middle English*, started from *Ḥ*, an average-sized letter which was able to be ready first.

The CAD was aptly described by Oppenheim (1966, 144) as a "utopian project" which "aims, on the semantic side, to relate meanings to the social context and technological background in which the references occur." For this reason, it is often mentioned that CAD is an encyclopedia more than a dictionary.

How does an article in the CAD look like? It can contain, at most, five sections:

1. The heading section lists the word's lemma, indicates which part of speech it is, gives its translation, historical and geographical distribution (e.g. OB indicates "Old Babylonian" attestations, NA the "Neo-Assyrian" ones etc.), provenience, morphology, possible ways of writing the word in Akkadian, and cross-references (e.g. an adjective is cross-referenced with a corresponding verb);
2. the lexical section, which offers explanations found in the ancient lexical texts, commentaries, grammatical texts et. sim.;
3. the semantic section, referred to as "the heart of the article," relates the meaning or possible meanings of the word;
4. the discussion, which elaborates the word considered in the format of a short essay; and
5. the bibliography.

See Figure 1 for an illustration of the article's disposition.

As vividly chronicled in Reiner's memoire, the history of CAD is shaped by the many and profound clashes of opposing visions for the project. Most notably, at the early stages, the visions clashed over the very basic questions of the dictionary's organization and purpose. Leo Oppenheim, one of the later chief editors of the project, summarized the differences in opinion between Ignace Gelb and Benno Landsberger in the following way:

> Landsberger demands with vigor that the argumentation should precede the enumeration of references. Gelb, however, wants first to

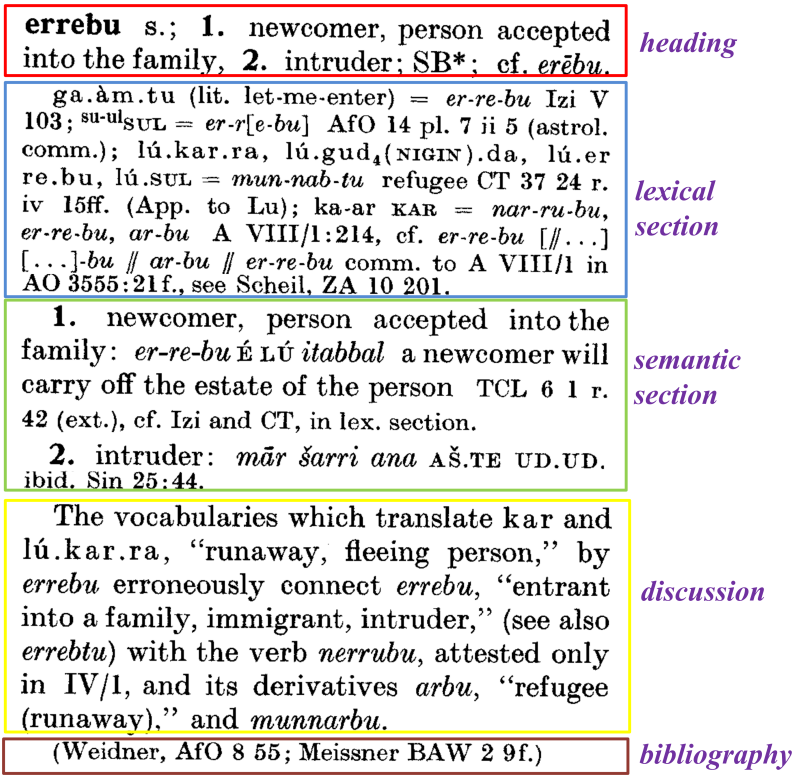

Figure 1. An example of a CAD article with all five sections (the example of the word *errebu* in volume E 1957).

> present the evidence with all references and then the discussion. It is rather obvious that both these "systems" reflect the individual psychological make-up of their originators; Landsberger prefers the dogmatic approach that is an adequate expression of his scholarly standing and temper, while Gelb wishes to follow the "objectivity" of the American linguistic school. (Leo Oppenheim, Contributions to the Discussions of the Standard Operating Procedure, memorandum of spring 1954; cited in (Reiner 2002, 24)).

Furthermore, for instance, in the early stages of the production, some scholars and most notably Gelb, insisted that the whole dictionary must

be more or less completed before even a single volume can be sent into print. The others, such as Oppenheim, insisted that the dictionary must keep producing volumes from the outset in order to keep the momentum, but also convince the academic sponsors that the valuable work is being done on the project, see (Reiner 2002, 15–16). Even later the relative speed with which the project developed concerned scholars who considered that not enough time was spent in examining the words thoroughly and with full view of the evidence – this might have been one of the reasons that caused the fiery departure of the famous assyriologist Thorkild Jakobsen from the project, see (52–62). Oppenheim's view that insisted on efficiency in the end prevailed – probably for the better, because, as we now know, one would have had to wait for a very long time before all the volumes could be considered prepared and it is unclear that the excitement about CAD, which attracted many external readers to help the project by writing articles, would have subsisted if no results at all were ever becoming available.

The separate publication of volumes – the publication dates of the earliest and the most recent ones are separated by more than half of a century – caused important differences between the volumes. It is well know, for example, that the first published volume, Ḫ, features a very different layout than the rest (see further in (38)). It is also interesting to note that, given that the dictionary cites copious passages relevant for the word in question, the translations of the same line of text, as they recur in different volumes as references for words beginning with different letters, often vary significantly, not only in style, but also in meaning. Observe the following example, citing three translations of the part of the line XI 131 of the *Gilgameš epic* in the three different volumes of the dictionary:

– Volume K (1971), s.v. kalû, meaning 5: "the sea grew quiet";
– Volume Š.3 (1992), s.v. šuḫarruru, meaning 2: "the sea grew calm";
– Volume T (2006), s.v. tâmtu, meaning 1a: "the sea subsided".

Even though it is clear that the sea calmed down, the exact nuance of the line in question is expressed differently in the three translations cited.

Furthermore, assyriology has made a lot of progress in the course of the half a century during which the publication of the CAD volumes slowly progressed. Newly discovered texts were edited and published, while the already familiar texts received more nuanced treatments in the light of the newly discovered evidence, which means that the citations which appeared in the earlier volumes of CAD are in some cases, in the light of the new editions, now outdated. It also occasionally happens that the one and the

same verse is registered differently in different volumes of CAD. For instance, the same verse is referred to as "En. el. I 142" in volume D (1958: 16) and as "En. el. I 143" in volume K (1969: 584).

Such minor inconsistencies and idiosyncrasies, while jarring, must be expected for projects the size and the scope of the CAD, the benefits of which outweigh its shortcomings by a wide margin.

3 Computers?

Some of the ambitious dictionary projects which I have mentioned in the last section have successfully embraced modern technologies, which facilitate the navigation and orientation in the mass of the gathered evidence. Take for instance the already mentioned and ongoing *Thesaurus Linguae Latinae* project, which covers the Latin texts up to 600 AD, the database of which covers all the lemmata from *a* to *resileo* (as of June 2023).[7]

It is also true that the field of Assyriology has become increasingly open to the novel and innovative approaches, not only in the domain of archaeology, where cutting-edge technologies are now able to reveal extraordinary data about the ancient objects, but also in the domain of textual criticism.[8] There have also been databases of cuneiform texts in transliteration and translation – the most important ones include: the *Cuneiform Digital Library Initiative* (CDLI), which seeks to collect and make available for free use roughly 500,000 cuneiform tablets and list them together with their photographs, physical information, relevant publication references, chronological information etc.;[9] the *Open Richly Annotated Cuneiform Corpus* (ORACC) which is developing a growing corpus of annotated cuneiform texts;[10] *Archives Babyloniennes* (ARCHIBAB), an electronic corpus of the Old Babylonian archives (https://www.archibab.fr/),[11] etc.

On the other hand, the CAD's relationship with computer technologies has been much less extensive. As early as in the 1960s, there have been attempts to convert the dictionary files into an electronic format. Reiner (2002, 78–79) gives an account of the meeting she and Oppenheim had with Robert

7. TLL - Thesaurus linguae Latinae

8. Gordin et al. (2020) presents a method for transliteration of cuneiform signs using Natural Language Processing (NLP).

9. The CDLI Collection

10. ORACC – Open Richly Annotated Cuneiform Corpus

11. ARCHIBAB – Archives Babyloniennes

Ashenhurst, a pioneering expert in computer sciences, where they learned that it would take up to three years to convert the file cards. Oppenheim's response "I don't have three years to waste" sealed that discussion. Nevertheless, in the mid 1960s a project was initiated to compile the reference index for each volume upon its appearance, but stopped after only the references for the volume "B" were completed, since the process proved to be too time-consuming, expensive and liable to errors.

CAD is currently available for download online, free of charge, from the website of the Chicago Assyrian Dictionary Project in a series of pdf documents. While this is undoubtedly useful, it is clear that navigating 10,000 pages of pdf documents (just as much as handling 10,000 pages of printed materials) occasions a lot of problems. The only way to search the documents is through the "find" function, which is often misleading and unhelpful. For instance, if one is looking for an article on a particular word, it is necessary to go through all the mentions of that word in the volume, because it may well be cited in many of the examples for the other words. If, on the other hand, one is looking for a particular reference (for example, "En. el. II 29"), the search can be offset by the fact that several lines are sometimes cited together and thus take shape unrecoverable by the "find" function (for example, "En. el. I 142, also II 29, III 33 and 91"). Furthermore, the search for certain references can be offset if they can be embedded into a larger framework of references – for instance, if one is looking for all the citations from the first book of *Enuma Eliš* ("En. el. I"), one will necessarily also get the references to the second, and the third book ("En. el. I*I*"; "En. el. I*II*"), etc.

It is clear that CAD could benefit from its conversion into a digital form, as has already been done for many historical dictionaries, e.g. *The Oxford English Dictionary* (Weiner 2008). Its transformation into a database, annotation according to some standard, like *Text Encoding Initiative*,[12] text parsing that can reveal inconsistencies and enable linkings at different levels, would facilitate the use of the dictionary to scholars and provide them with many new ways to use it.

4 Final remarks

It is easily seen that the *Chicago Akkadian Dictionary* is an essential tool in the discipline of assyriology. Assyriologists use it when they encounter

12. TEI – Text Encoding Initiative

new and as yet undeciphered texts (which happens all the times as new tablets are uncovered), and equally so when they return to the classical texts, which have been known for more than a century, such as the *Gilgameš* poem, whenever they seek to improve their understanding of a word in a particularly difficult passage. As we have seen, the project such as CAD has never been self-evident, and its path has never been smooth: the fact that it has lived to see the light of day in the full is the reason enough to celebrate the tremendous achievement and the assiduity of the generations of brilliant scholars who worked on it.

But I would like to end with a question: can we build on their effort to achieve even more? Can information technologies, which are experiencing their heyday, and which are, as we have seen, used amply and to the great profit in the field of assyriology, bring the much needed modernization of the CAD project by improving its searchability, the interconnectedness of its references and the uniformity of its citations? The suitable application of NLP methods might enable the users of the dictionary to search, for example, a reference to a particular verse or, rather, a particular book of a particular poem throughout the corpus, and allow for a search that prioritizes attestations of a particular period or provenance, and, finally, they might replace the toil of navigating 26 long and cumbersome pdf documents by a simple electronic search tool. It remains to be seen, but in a world that is being shaped and constantly changed by the innovative technologies, it is to be hoped that its zeal will once again put wind in the sails of the inimitable CAD project.

References

Daniels, Peter T. 2020. "The Decipherment of Ancient Near Eastern Languages." In *A Companion to Ancient Near Eastern Languages,* edited by R. Hasselbach-Andee, 3–26. John Wiley & Sons, Inc.

Gelb, Ignace J. 1964. "Introduction to the volume A/1." In *The Chicago Akkadian Dictionary A/1.* Glückstadt: J. J. Augustin.

Geller, Markham J. 1997. "The Last Wedge." *Zeitschrift für Assyriologie* 87:43–95.

Gordin, Shai, Gai Gutherz, Ariel Elazary, Avital Romach, Enrique Jiménez, Jonathan Berant, and Yoram Cohen. 2020. "Reading Akkadian cuneiform using natural language processing." *PloS one* 15 (10).

Oppenheim, Leon. 1966. "Akkadisches Handwörterbuch unter Benützung des lexikalischen Nachlasses von Bruno Meissner (1868-1947), bearbeitet von Wolfram von Soden." *Journal of Near Eastern Studies* 25 (2): 129–155.

Reiner, Erika. 2002. "An Adventure of a Great Dimension: The Launching of the Chicago Akkadian Dictionary." *Transactions of the American Philosophical Society* 92 (3).

Rojas, Felipe. 2023. "Urartian Stelae in Late Antique and Early Medieval Armenia." *Iran and the Caucasus* 27 (2): 129–155.

Roth, Martha T. 2010. "How We Wrote the Chicago Assyrian Dictionary." *Journal of Near Eastern Studies* 69 (1): 1–20.

Stolper, Matthew W. 1991. "The Chicago Assyrian Dictionary at Seventy." *Oriental Institute News & Notes* 129.

Weiner, Edmund. 2008. "The electronic OED: The computerization of a historical dictionary." In *The Oxford History of Lexicography,* edited by A. P. Cowie, 378–409. OUP Oxford.