CHAPTER 18

Fear in Akkadian Texts: New Digital Perspectives on Lexical Semantics

Saana Svärd, Tero Alstola, Heidi Jauhiainen, Aleksi Sahala and Krister Lindén

1 Introduction*

The study of emotions in ancient Mesopotamia is still very much in its infancy. There has been some scholarly interest regarding Mesopotamian emotions, but before this volume much of it has been conducted from a medical or psychological perspective.1

Understanding emotional practices across such a great geographical, cultural and temporal range is fraught with complications. An important topic that has been hardly touched in Mesopotamian studies yet is the concept of “emotion” itself. The concept of “emotion” is not universal, and in Akkadian there is no term that could serve as a translation of it or other such categories (e.g., “affect,” “feeling”). From a comparative perspective, it has been demonstrated that in the Hebrew Bible, lexemes expressing explicit emotions are scarce and other linguistic expressions, such as literary devices, bodily sensations, certain actions or clusters of actions, are used to indicate emotions. Additionally, emotional responses often refer to social situations, suggesting that the social dimension of emotions was more important than the internal feelings of an individual.2 All these observations should be examined in Akka-

* Work on this article was jointly conducted by all the authors, but for the most part the division of work was as follows: Svärd wrote Sections 1, 2 and 6, Jauhiainen 3 and 4.2, Sahala 4.1, and Alstola, Jauhiainen and Svärd 4.3–4. Svärd and Alstola analyzed the results and wrote Section 5. Jauhiainen processed the dataset. Sahala designed the weighting algorithm for PMI and wrote the tool for calculating the PMI scores. Alstola coordinated the design for the workflow (Section 4.4). Lindén directed the language technological work and was crucial for the success of this cross-disciplinary project. A technical note: in this article we follow the volume convention of using ḫ instead of h, but in our data we use the regular h. The Akkadian spellings of words and their translations follow the lemmas in Oracc.

1 For a more comprehensive overview of the study of emotions and Mesopotamia, see the Introduction in this volume.

2 Mirguet 2016.
dian as well as in future research. For this article, however, we merely examine certain lexemes relating to “fear” and make no claims to have covered everything related to “fear” in Akkadian language with this inquiry.

The field of the history of emotions has bloomed in the past few decades. The question of the possible universal nature of emotions has been at the heart of many discussions in this field of study. Emotions vary greatly depending on time, place and context, but how much and why? Are some emotions universal and rooted in the human mind and body? A related issue is the methodological process itself. How can we access the emotional lives of individuals in the past? What is the relationship between the emotional expressions in our text material and the actual lived and embodied emotional states?

In linguistics, universalism has also been a debated topic. Some scholars have argued that individual languages are mere variations of the underlying universal cognitive language model. On the other hand, linguistic relativists proved that native speakers of different languages do have real differences at the conceptual level. This has implications for understanding words related to emotions as well—how do emotional concepts of another language differ from the concepts in “our” language? We do not disregard the possibility of some universal emotions, but for this article we concentrate on examining the particularities of the Akkadian language. Our departure point is that linguistic expressions both influence how emotions are felt and allow individuals to express emotions.

Thus, we study what kind of conceptual maps might emerge for Akkadian emotions. In this article, we examine the semantic lexical field of five verbs that are commonly translated as “to fear” or similar in English, namely, adāru, galātu, palāḫu, parādu and šaḫātu. In addition to the verbs, we also analyze nouns and adjectives derived from them. This selection of lexemes is already a methodological choice, as we do not know if these words formed any kind of recognizable map of “fear” as we understand the emotion in modern English. Nonetheless, these verbs form an entry point by which we can at least attempt to understand how the native Akkadian users themselves may have perceived the relationships between these five lexemes and their relationships to other words.

Methodologically, we hold that the meaning of any given word, including words that express emotion, lies in its relationship with other words. A word is

---

3 For a good introduction to these questions, see, for example, Plamper 2015 and Eustace et al. 2012.
always a member of a larger group of words and shares certain aspects of meaning with them. The methods of language technology help us trace semantically relevant properties of the lexemes of interest in large corpora, and this statistical analysis provides tools for Assyriologists to reflect on the semantic domains of specific words.

Lexical semantic work on emotions in Mesopotamia is scarce. In addition to the standard dictionaries and the articles in this volume (see especially the excellent article by Steinert on fear and other emotion words), there are only a few studies that focus on lexical semantics instead of discussing how affect and emotions are portrayed in Sumerian and Akkadian texts on a more general level. The most important research on lexical semantics includes Jaques’s work on Sumerian emotion words and some of their Akkadian equivalents, Rendu Loisel’s dissertation on sound and emotion, and Wasserman’s discussion of the semantic field of love. Another perspective is a comparative study across Akkadian and West Semitic languages (for example, the new volume on hate by Riley). In this article, by conducting statistical analysis with language technological methods, we aim to advance lexical semantic work on Akkadian fear words, especially for lexemes that occur frequently.

The article is structured as follows. Section 2 discusses previous lexical research on “fear” in Akkadian, based on Akkadian dictionaries. Section 3 describes our dataset. The Akkadian text data we are using comes from the Open Richly Annotated Cuneiform Corpus (Oracc). We are relying on Oracc annotations to a large degree, but we have used some preprocessing to enrich the data set. Section 4 describes our methodology very briefly, with references to our previous work, where more details can be found. In sum, we use two computational methods—Pointwise Mutual Information and fastText—to study the semantic contexts of the chosen emotion words. Furthermore, network analysis is used to explore and visualize the relationships between the words. Combined use of these methods and analysis of some lexemes in their original word context yield data-driven results, which are examined in Section 5. The article is concluded in Section 6 with a general discussion of the results. By analyzing and highlighting the similarities and differences in the semantic fields of these five fear verbs, we create for the first time a semantic map based on statistical analysis of how these verbs relate to

4 Jaques 2006; Rendu Loisel 2010; Wasserman 2016, 32–44, 52–62.
5 Riley 2017.
6 Svärd et al. 2018; Alstola et al. 2019a, 2019b, in preparation; Jauhiainen and Alstola forthcoming.
each other and other lexemes in our corpus. Our research data and network visualizations are openly available online.\textsuperscript{7}

2 Previous Lexical Research on Fear Words in Akkadian

We selected five common fear-related verbs for analysis on the basis of translations given in the standard dictionaries: namely, \textit{adāru}, \textit{galātu}, \textit{palāḫu}, \textit{parādu} and \textit{šaḫātu}. We disregarded the verbs that appear rarely in our corpus. As our main source material for the previous lexical research on these words, we used the standard dictionaries of the field: \textit{CAD}, \textit{CDA} and \textit{AHw}. We also checked the online Supplement to the Akkadian Dictionaries (eSAD) for occurrences, but found nothing that would add to our analysis of \textit{CAD}.\textsuperscript{8}

The verb \textit{palāḫu} is widely attested in Akkadian sources and can be used without an object to represent a general state of fear, or objects and prepositional phrases can be used to indicate reasons for the fear. The reasons are manifold and the verb is used ubiquitously (i.e., not occurring solely in a certain text genre, formula, time period or location) to express fear in many different contexts. The verb also has a more general feeling of unease connected to it, and it can be translated into English as the state of being worried or anxious.

The subject of the verb can be an individual person who is afraid or the whole region, city or people. Gods can also be afraid, albeit rarely. Martial contexts do appear frequently, as \textit{palāḫu} relates to the fear that the enemy will attack. Quite often the gods or something unearthly (like \textit{melemmu} “radiance”) inspire fear and awe in people. In these cases (and others as well), fear should not be seen as a negative emotion but a correct and appropriate response to the presence of the numinous. This positive aspect of \textit{palāḫu} is shown in such \textit{CAD} translations as “to be reverent, respectful” and “to respect, honor and venerate.” The object of reverence is almost always the king or gods, or something connected to them (e.g., the gods’ word/command, an oath or the palace). Without an object it can indicate a general state of good manners and proper conduct.\textsuperscript{9}

\textsuperscript{7} http://doi.org/10.5281/zenodo.3634325 (February 4, 2020).
\textsuperscript{8} Available at http://altorient.gko.uni-leipzig.de/etymd.html (January 10, 2020).
\textsuperscript{9} This was also pointed out by Gruber (1990, 420), who writes about the meaning of \textit{palāḫu}: “At the center of the spectrum is worry about impending danger. At the one end of the spectrum is the tendency to respond to the impending danger by adopting a positive course of action.” Such a positive course can be translated into English as “being reverent towards an authority (e.g., a god).” He further writes: “At the other end of the spectrum is pathological anxiety, which leads to total inaction.”
**Table 18.1** The translations of five fear verbs in the standard dictionaries of Akkadian

| Word   | Translation in CAD | Translation in CDA | Translation in AHw                      |
|--------|--------------------|--------------------|----------------------------------------|
| palāḫu | to be afraid, worried, respectful; to serve | to fear, to revere | (sich) fürchten, verehren               |
| šaḫātu | to fear, to respect | to be afraid, to fear, to hold in awe | (sich) scheuen, fürchten                |
| adāru  | to fear and respect, to be afraid | to be afraid of, to fear | finster sein, (sich) fürchten         |
| galātu | to twitch, to quiver, to ejaculate, to be nervous, to fear | to tremble, to be afraid | erschrecken, zittern                   |
| parādu | to be fearful, disturbed, restless, upset, confused | to be scared, terrified | erschrecken                            |

When “parents” are the object of *palāḫu*, the translation can be “to treat respectfully” and “to serve, to care for.” The latter has perhaps led to a more specific meaning; in some economic contexts (first millennium BCE), *palāḫu* means “to fulfill a financial obligation by working.”

Bodily attributes are connected to *palāḫu* as well as many other emotions: the heart (*libbu*) is mentioned as a common bodily location for fear, and reverence can take a physical form as well (for example, the “lifting of arms” mentioned in a Neo-Assyrian letter).

The verb *šaḫātu* “to fear, to respect” perhaps occurs in more secular contexts than the other fear verbs, although in the *Erra Epic* even gods can be afraid and the word can mean “to respect, fear the gods.” Martial contexts appear as well, but the verb is used in various situations (also related to medical conditions) and it is difficult to pinpoint distinct contexts. AHw gives the word a distinct translation of “(sich) scheuen, fürchten,” suggesting that in addition to fear, it meant “to shy away,” indicating a general nervousness. The verb *šaḫātu* occurs in letters, literature and royal inscriptions over a long period of time. The word has two well-attested homonyms (and a rare homonymous noun, not included here): *šaḫātu* A (noun) “side, inside corner” and *šaḫātu* B “to wash, to clear.”

---

10 This is the meaning that is transferred to later Semitic languages as well. In Mishnaic Hebrew and Aramaic, the verb can mean “to serve, work” and in Arabic “to till the land” (AHw).

11 SAA10 198: “And may I, blessing the king, my lord, stand before the king, my lord, and revere (him) wholeheartedly with my arms (lifted in the gesture of blessing); and when my arms become weary, may I use up my strength with my elbows” (in Akk. *ina a-ḫi-ia la-ap-lāh, SAAo*).
The latter word in turn has a connection with šaḥāṭu “to strip off, remove” (see p. 86 in CAD Š/1), and it is often difficult to make a distinction between the two verbs.

CAD treats adāru as three homonymous words: adāru (a tree), adāru A “to be worried, to become obscured” and adāru B “to fear, to respect.” The semantic contexts listed for adāru B in CAD are fairly straightforward. Humans are mostly subjects, while gods, oaths and parents are common objects of the verb. For oaths, gods and parents, there seems to be a connotation of positive awe and respect. There is also a more negative side to the verb: it can mean “to be afraid of someone or something” (for example, the elements, like waves or snow). Adāru commonly has clear martial and battle connotations, and hence it relates to kings as well. The verb also occurs together with palāḥu. All these connotations are evident in AHw, which suggests that adāru A and B (listed separately in CAD) are in fact the same verb, with connected meanings of “fins- ter sein” / “going dark” (e.g., heavenly bodies, gods, a sick person or face) and “(sich) fürchten” / “to fear.” The question of adāru A and B is complicated, but there is no need for extensive details here.12 For the purposes of this article, we are following CAD and eSAD and are only interested in adāru B “to fear.”

As for galātu, the contexts are similar in CAD and AHw. In both, “to twitch,” “to tremble,” “to shake” and so forth form the root meaning. In general, there are clearly negative connotations of some kind of movement. This “twitching” can relate to body parts (which explains the CAD translation of “to have a premature emission”) or the whole body, especially in the case of being startled.

12 Von Soden simply notes in AHw: “trotz der verschiedenen Bedeutungen wohl dasselbe Verbun” (“despite the different meanings probably the same verb”). There are grammatical grounds for separation, but CAD admits to a certain degree of confusion between adāru A and B (A/1, pp. 107–108), suggesting, “The two verbs ... have been separated here primarily on semantic grounds ... adāru B ... for which the semantic range of religious awe, respect, and service has been assumed in contrast to the connotation of excitement, unrest (psychological and physiological) which seems to be characteristic of adāru A.” This is, however, not a clear-cut difference, as the semantic connotation of unrest, which CAD connects to adāru A, can also translate to fear and apprehension. In CAD, adāru A is primarily connected with heavenly objects that can become unclear, obscured, eclipsed or disturbed in some other way, so that one cannot see them properly. Based on this, adāru could be imagined as a verb of “being obscure” or “disturbance in the natural order of things.” For humans, the face is the object that becomes obscured. It is difficult to pinpoint an English translation for the concept of “obscured/disturbed person” in the Akkadian language. Therefore, the translations in CAD vary between emotional responses of having fear or grief or being confused, sad, troubled, worried, apprehensive, disturbed, annoyed or impatient. Thus, fear is among the sentiments expressed with adāru A, although it is not a major one. See also Jaques 2006, 181–184.
or having fear. A related meaning is “a person becoming restless or nervous” (also the sea can be restless). Quite often this twitching or restlessness happens either during sleep or illness. The meaning “to be or become frightened” obviously has its roots in the bodily reaction of “trembling with fear.” Fear can be connected with the heart (*libbu*), and the subjects can be humans but also gods and troops. Fear can be inspired by demons, animals, ghosts, monsters, the sea, *melemmu*, divine statues and places. It seems that often animals can be frightened but very rarely gods.

The meaning of the verb *parādu* seems to be connected with the idea “to disturb in some way the correct state of things.” The word often appears with *libbu* in the sense of a “disturbed heart,” which can refer to being fearful or restless and worried. Lands, kings, troops and even gods and benevolent spirits can be afraid. A cause of fear can be, for example, dreams, an enemy, a snake or a ghost. The verb seems to form an antithetical pair with *nuāḫu* (“to rest”), being connected with other verbs denoting movement (e.g., *dalāḫu* “to disturb”). It also occurs (often with *libbu*) as a sign of a medical condition or an omen. In divination, the verb can refer to confused signs or omens. In Old Assyrian, it appears in an interesting idiom “the market is restless.”

### 3 Dataset

#### 3.1 Data from Oracc

Our data comes from the Open Richly Annotated Cuneiform Corpus (Oracc), one of the largest, and regularly updated, digital corpora of Sumerian and Akkadian texts. Our data is a snapshot of the corpus from February 2019 and consists of 16,487 texts and over 1.8 million words.

The text material used for this article contains all texts tagged in Oracc as being written in the Akkadian language. Some of the documents in Oracc are multilingual, but we did not include such texts (e.g., those written in Sumerian in addition to Akkadian). We did, however, take into account texts where

---

13 The research data used for and generated during our research is available online at Zenodo: http://doi.org/10.5281/zenodo.3634325 (February 4, 2020).
14 http://oracc.museum.upenn.edu (February 5, 2020).
15 Because Oracc subprojects designated with the words “Scores” and “Sources” only contain duplicate texts, which are available elsewhere in the parent projects, we decided to leave them outside our data.
16 Some of the texts in the SAAo project have been tagged as being written in either Neo-Assyrian or Neo-Babylonian. These texts were included in the dataset.
the language of the whole document had not been indicated, but where all the individual words were tagged as Akkadian. Thus, we rely solely on the metadata provided for each document in Oracc. The texts tagged as being “lexical” in Oracc were omitted, since they refer to a great number of rare words, many of which are only attested in the lexical lists. We also searched for duplicate texts in the corpus and have kept only one copy of each for the analysis. The dataset used for analysis consists of 7,346 texts, most of which are from the Neo-Assyrian period (5,638) and come from Nineveh (4,641). Many of our texts are various kinds of letters (2,247), but royal inscriptions (1,487) and legal transactions (1,434) are also well represented, as well as astrological and astronomical texts (613). The texts contain over one million words.

The data for this article is part of a larger dataset that concentrates on research on “love” and “anger” in addition to “fear.” Therefore, the words of interest in the dataset include—in addition to five verbs of fear—the verb rāmu (“to love”) and eight verbs related to anger: agāgu, ezēzu, kamālu, labābu, raʾābu, šabāsu, šamāru and zenû. The inclusion of these verbs in our dataset does not have a direct effect on our analysis of the five fear-related verbs.

3.2 Preprocessing of the Data

Since Akkadian is an inflected language, we decided to use only the lemmas—the dictionary or citation forms—of the words. We wanted to explore the effects of preprocessing of linguistic data on the results. In order to study these, we made two different versions of our dataset to be analyzed with two language technological methods: Pointwise Mutual Information (PMI) and fastText (see Section 4).

Level 1. Basic. In the first level, we use lemmatized texts downloaded from Oracc with some minor changes. As annotation of a text always involves interpretation by a scholar and because Oracc is composed of a number of sub-projects, the way a word is annotated in different texts is not always consistent. Therefore, we have standardized divine names and some place names,

---

17 A list of duplicate texts is found in the textsInformation/duplicateTexts file at Zenodo. On our criteria for working with duplicates, see Alstola et al. 2019a, 163.
18 A list of genres in our data is in the textInformation/genres file at Zenodo.
19 A list of texts used in our analysis is available at Zenodo in the “textInformation” (textInformation/documents_used) folder, whereas the analyzed texts themselves can be found in the “textsAnalysed” folder.
20 We only used verbs, nouns and adjectives for the analysis; the rest of the words were replaced with an underscore as a placeholder.
and combined different personal and royal names, as well as some other word classes, into one lemma each (e.g., “PERSON” or “ROYALTY”). We analyzed the above-mentioned verbs of fear, love and anger and their derivatives when they appeared more than nine times in our dataset. This threshold was chosen so that each verb of emotion or at least one of its derivatives is included in our analysis. Accordingly, some rare verbs of anger like kamâlu are not included in the analysis but represented by their more common derivatives (e.g., kimîltu “wrath”). The lemmas (with their respective number of attestations) denoting fear were adāru (139), ādiru (25), adirtu (13), adru (177), šutadduru (12), galâtu (35), gilîttu (35), palâhu (394), palîhu (83), pulîhu (58), puluštîtu (101), parâdu (114), pardu (17), pîrîttu (38), šaḥâtu (47) and šaḥtu (31).

**Level 2. Roots combined and homonyms separated.** In the second stage, we did two separate processes. First, we separated homonymous words from each other on the basis of translations provided by the Oracc metadata. For instance, rabû with the translation “be(come) big” is not the same word as rabû “set.” We used automated steps to keep words with synonymous but slightly different translations together. The occurrences of the word rabû with the translations “be(come) big” and “to be big” can be regarded as one word and not considered homonyms. Moreover, two words are regarded as one word if they belong to different parts of speech but derive from the same verbal root, have an identical base form and have an identical English translation in Oracc after the translations are harmonized. For example, the different translations of the verb rabû (“to be / become big”) were harmonized as “big,” which led to the adjective rabû being considered the same word as the verb. Some synonymous translations (for example, rabû “big” and “great”) are difficult to detect automatically. Thus, we manually improved the results of the automatic separation of homonyms. We compiled a list of all words that had two or more different translations after the automated separation. Based on this list we manually reduced 2,350 words with synonymous translations to 1,114 words. In the texts to be ana-

---

21 For further details on how we processed the data, see Svärd et al. 2018, 227–228; Alstola et al. 2019a, 162–164. The lists of standardized divine and place names and information on the standardization process are available in the Zenodo files lists/gods_combined_variants and lists/places_combined_variants.

22 A list of all words after we had automatically separated homonyms (and combined the derivatives to their verbal roots) are in the lists/words_automaticHomonymSeparation file at Zenodo.

23 The list of Akkadian words with the same meaning but different English translations (e.g., palâhu “to fear” and “to be afraid”) is available in lists/notHomonyms at Zenodo. We thank Aino Haavisto for her help in compiling this list.
lyzed, homonyms are distinguished by a running number starting with “_1.” Even if a word has no homonyms, it is marked with “_1” placed after the word.

Secondly, in order to have more data per word for our methods to process, we combined all the derivatives of the emotion verbs with their verbal roots.\(^{24}\) In level 1, all homonymous words (such as the word *adru* with the meanings “be afraid,” “dark” and “threshing floor”) were in the analysis considered as just one word. As the homonym separation was done first, only the instances of *adru* that in Oracc have the translation “be afraid” were combined with *adāru*_2 “be afraid” in level 2.\(^{25}\)

After the process we had, for example, three tokens for the lemma *rabû*: *rabû*_1 “big” (4,299 attestations), *rabû*_2 “set” (184) and *rabû*_3 “fourth” (1). We did not standardize the spelling of lemmas, and thus *rabû*_3 “fourth” and *rebû*_1 “fourth” are kept separate. The tokens of the verbs of fear are *adāru*_2 (159 attestations), *galātu*_1 (87), *palāḫu*_1 (637), *parādu*_1 (169) and *šaḫātu*_2 (51). The tokens of the verbs of love and anger are *rāmu*_1 (488), *agāgu*_1 (69), *ezēzu*_1 (205), *kamālu*_1 (29), *labābu*_1 (22), *raʾābu*_1 (55), *šabāsu*_1 (42), *šamāru*_1 (65) and *zenû*_1 (72). Only *rāmu*, *adāru*, *šaḫātu* and *šamāru* had homonyms.

The rich dataset described in this section is fully available online (http://doi.org/10.5281/zenodo.3634325), with readme files and explanations. The dataset includes seven folders: 1) “lists” includes, for example, lists of standardized divine names and place names as well as frequency lists of the words analyzed, 2) “network-data” includes Gephi datasets and the raw data used to create the graphs, 3) “oracc_data” contains all the words extracted from Oracc in February 2019, 4) “scripts” includes the scripts created to analyze our data with PMI, 5) “textInformation” contains metadata on the whole text corpus we analyzed, 6) “textsAnalysed” includes the actual text files that were analyzed with PMI and fastText, and 7) “visualizations” contains visualizations of the networks as PDF files.

4 Methods

4.1 *Pointwise Mutual Information*

Pointwise Mutual Information (PMI) is one of the most popular statistical association measures used in automatic collocation extraction. The core idea of

\(^{24}\) We thank Jason Silverman for suggesting this method to us. A similar approach is used by Andrew J. Riley in his recent study on hate in the ancient Near East (Riley 2017, 11).

\(^{25}\) The list of the words after all the preprocessing was done is in the *lists/wordsWithTranslations_level2* file at Zenodo.
PMI is to describe word association as a logarithmic ratio of the observed joint probability $p(a,b)$ of the words to the expected chance of their co-occurrence under the assumption of probabilistic independence $p(a)p(b)$. In other words, PMI measures the ratio of the observed word co-occurrence probabilities to their co-occurrence probabilities in a hypothetical scenario where the word order of the corpus was randomized and all syntactic and semantic information was lost.

$$PMI(a; b) = \log_2 \frac{p(a,b)}{p(a)p(b)}$$

PMI excels in discovering syntagmatic semantic relationships. The characteristics of these relationships depend on the distance of the words from each other, of which the maximum value can be controlled by a parameter called a window. Small windows (for example, the distance of three words on either side of the word of interest) are best for finding compound words, multi-word expressions and idioms, whereas larger windows capture more abstract semantic dependencies (for instance, specific terminology related to a certain topic). For studying the emotion words, we used a large symmetrical window of ten words on either side.

As PMI has a well-known tendency to favor low-frequency words, and our corpus was too small to define adequate frequency thresholds without losing too much data, we chose to use an improved variant of PMI called PMI$^2$. This measure is identical to PMI, with the exception that $p(a,b)$ is squared in order to reduce the low-frequency bias. To make the association scores more intuitive and interpretable in the visualization tool Gephi, we removed the logarithm to flip the negatively oriented PMI$^2$ scores onto the positive axis. Positive PMI$^2$ is defined as:

$$PPMI^2(a; b) = \frac{p(a,b)^2}{p(a)p(b)}$$

When analyzing Akkadian texts, PMI results are often distorted by repetitive and heavily formulaic content in the corpus. For example, a single word may have several very strong collocates, which all emerge from slightly different

26 Church and Hanks 1990.
27 Manning and Schütze 1999. For syntagmatic and paradigmatic semantic relations, see Svärd et al. 2018, 230–232.
28 Daille 1994.
29 This corresponds to $PPMI = 2^{PMI^2}$ as defined in Role and Nadif 2011, which realigns the lower bound of PMI$^2$ scores from $-\infty$ to 0 and the upper bound from 0 to 1.
copies of the same text. In order to reduce the significance of such repeating content and to highlight more interesting and free use of language, we developed a method for weighting the PMI scores based on their contextual similarity. The weighting algorithm scans every word position in every collocational window for every pair of words, and calculates the relative frequency of unique context words in that position. By context word, we mean any word that is neither of the words we are calculating the association score for. For example, let us assume that we are interested in the association between words A and B, which co-occur three times in a forward-looking window of size 5. We represent the context words found in our windows by the symbols \( w, x, y \) and \( z \).

| Word position in the window | 1   | 2   | 3   | 4   | 5   |
|-----------------------------|-----|-----|-----|-----|-----|
| Co-occurrence 1             | A   | x   | y   | z   | B   |
| Co-occurrence 2             | A   | x   | y   | z   | B   |
| Co-occurrence 3             | A   | w   | z   | B   | x   |
| Rel. freq. of unique context words | \( \frac{2}{3} \) | \( \frac{2}{3} \) | \( \frac{1}{2} \) | 1 |

We calculate the context similarity factor \( F \) by taking the average of the relative frequencies of unique context words. We ignore the keyword A and its potential collocate B from the counts to prevent all contexts from being penalized. Formally this can be expressed as

\[
F = \frac{1}{w-n} \sum_{i=1}^{w-n} f_i
\]

where \( f \) is the relative frequency of unique context words in position \( i \), \( w \) is the window size and \( n \) is the number of positions completely occupied by non-context words (e.g., A in position 1 in our example). In the case of our example above, we calculate \( F \) as \( \frac{1}{5-1} \left( \frac{2}{3} + \frac{2}{3} + \frac{1}{2} + \frac{1}{1} \right) \approx 0.71 \) and apply it to our association score as \( \text{PPMI}^2(A;B) \times F \).\(^{30}\)

In this article, PMI is used to find words (collocates) which co-occur with the emotion words under scrutiny. While these collocates are not necessarily words of emotion, they are likely to highlight the contexts in which our words of interest occur. For example, if a verb of fear is commonly used to denote fear

\(^{30}\) The factor can be used with other PMI variants as well and may be applied either before or after taking the logarithm, depending on how strong similarity filtering is desired. The factor removes positive association in cases where most of the co-occurrences appear in very similar contexts. However, with PMI\(^2\) and PPMI\(^2\) multiplying the final score is convenient, as both measures sort the scores according to their distance to the perfect dependence instead of independence.
of enemy troops, words such as “soldier” and “battle” might be attested on the list of the closest PMI collocates.

4.2 FastText

Word vectors are numerical vector space representations of the word forms of a text or set of texts.\textsuperscript{31} Such word vectors model semantic and syntactic regularities between words. Words can be used in various contexts which bring different images to mind,\textsuperscript{32} or as Firth put it already in 1957: “You shall know a word by the company it keeps.”\textsuperscript{33} Vectors of words appearing in comparable linguistic contexts have similar coordinates and cluster near each other in the vector space. FastText and word2vec are methods that use an artificial neural network model to generate such word vectors, and they have been shown to be better at finding semantic similarities of words than methods that count word co-occurrences.\textsuperscript{34}

Word2vec\textsuperscript{35} has two alternative models one can choose between: the \textit{Continuous Skip-gram} model and the \textit{Continuous Bag-of-Words} model.\textsuperscript{36} When encountering a word, the Continuous Skip-gram model tries to predict the words that surround the target word, whereas the Continuous Bag-of-Words model tries to predict the word itself from the surrounding words.

FastText is an extension of word2vec.\textsuperscript{37} In addition to words, fastText takes subword information into account by representing a word as smaller sequences of characters derived from that word. The position of a word in the vector space is determined by adding up the information on all the different sequences of characters. According to the developers of fastText, the Continuous Skip-gram model works better in fastText than the Continuous Bag-of-Words model,\textsuperscript{38} and that is also our observation. Therefore, our results have been produced using the Continuous Skip-gram model.\textsuperscript{39}

---

\textsuperscript{31} E.g., Pennington et al. 2014.
\textsuperscript{32} de Blois 2008.
\textsuperscript{33} Firth 1957, 11.
\textsuperscript{34} Baroni et al. 2014.
\textsuperscript{35} Mikolov et al. 2013a; Mikolov et al. 2013b; see also Svärd et al. 2018.
\textsuperscript{36} Mikolov et al. 2013b.
\textsuperscript{37} Bojanowski et al. 2017.
\textsuperscript{38} https://fasttext.cc/docs/en/unsupervised-tutorial.html (January 20, 2020).
\textsuperscript{39} As options to start fastText, we use negative samples 70 (instead of 5), sampling threshold 0.01 (0.0001), window size 10 (5), vector dimension 50 (100), and number of epochs 1,000 (5). For all other parameters we use the default values; see https://fasttext.cc/docs/en/options.html (January 21, 2020).
We use fastText to detect words which are most similar to the emotion words we are studying: the similarity of two word vectors indicates that the words in question have a similar semantic function in the language and can appear in similar contexts. If the corpus is large enough, the most similar word vectors are synonymous, antithetical or their meaning is very close to each other in some other way. When studying words of fear, one expects to find other fear words among the most similar words detected by fastText.

4.3 Network Analysis and Visualization

For visualizing our results we built networks (graphs) using Gephi.\(^{40}\) In such networks, words are represented by nodes and connections between words by edges (ties). One set of graphs was produced using the PMI results and another set using the fastText results. The edges in our graphs mark the similarity of words according to fastText and PMI. In the PMI graph, the edge weight between two words reflects their PMI score. In the fastText graph, the edge weight corresponds to the cosine similarity between two word vectors in the vector space. The size of a node in the graph indicates its degree, which is the number of edges connected to that node.

We chose to use a graph-drawing algorithm called ForceAtlas2 because its results have been shown to be the most interpretable in our previous research.\(^{41}\) ForceAtlas2 is a force-directed algorithm, which draws graphs by trying to push the nodes as far away from one another as possible while the edges try to keep connected nodes close to each other.\(^{42}\) The smaller the number of edges between two nodes, the stronger they are pulled together.

A subgroup within a network is called a community. In a community, nodes have many ties between one another but fewer ties to the rest of the network. Force-directed graph-drawing algorithms help visualize communities as grouped nodes within the network.\(^{43}\) We used the ForceAtlas2 algorithm with the so-called LinLog mode, making the communities in a graph tighter.\(^{44}\) In order to visualize the community structure more clearly, we used the built-in modularity algorithm in Gephi, assigning each node to a community.\(^{45}\) We then highlighted the communities in our graph by giving a different color to each of them.

\(^{40}\) Gephi is open source, free software and can be downloaded from https://gephi.org (February 4, 2020). See Bastian et al. 2009.

\(^{41}\) Alstola et al. 2019a.

\(^{42}\) Jacomy et al. 2014.

\(^{43}\) Noack 2009.

\(^{44}\) Jacomy et al. 2014.

\(^{45}\) The algorithm uses the Louvain method to assign nodes to communities. See Blondel et al. 2008.
4.4 Combining the Methods

Up to now we have described our methods individually. To summarize the sequence of analysis, (1) we built a file with each document of the dataset as one line. This was done separately for both levels of preprocessing. (2) Analyzing the two levels separately, we extracted collocations with PMI for all the emotion words we were interested in and (3) built word vectors with fastText and extracted the closest words for the emotion words under scrutiny. (4) We visualized the emotion words and their collocates and semantically similar words as networks with Gephi. We built graphs for the PMI and fastText results with both the 10 and the 50 closest words. The larger graphs of 50 also contain the 10 closest words, of course, but with additional words we could get an idea of the wider contextual domain of the words of interest. In the smaller graphs of the 10 closest words for each emotion word, it was easier to spot the links between words and, in some cases, the common contexts as well.46 (5) The analysis of these networks (presented in Section 5) moved in a hermeneutic circle. After examining the graphs, certain ideas of the context emerged. These different possibilities were examined by going back to the context of individual words.

For examining individual words and their contexts, we used the corpus search tool Korp, an online service provided by the Language Bank of Finland, which contains several versions of Oracc data.47 At the time of writing, the Korp version of Oracc that can be searched in the user interface is “May 2019.”48 The Akkadian text data in that version is very close to the dataset that we used in this article (see Section 3 above).49 Therefore, in what follows, “Korp” refers to the dataset “Oracc in Korp May 2019,” not directly to our research data.

The results of a query in Korp are presented as concordances (i.e., every instance of the word matching the query is listed line by line with all its neighbors). Furthermore, the matching words are highlighted in the middle of the display. Korp also allows for gathering statistics regarding the genre (or other

46 All of the graphs and the data used for creating them are available in the online repository Zenodo: http://doi.org/10.5281/zenodo.3634325 (February 4, 2020). In our graphs, the size of a node indicates its degree (i.e., the number of nodes to which it is directly connected). The colors of the nodes show communities assigned by the modularity algorithm in Gephi.
47 https://www.kielipankki.fi/support/korp/ (February 4, 2020); https://www.kielipankki.fi/corpora/oracc (February 4, 2020). Earlier versions of the data can be downloaded from the service.
48 Jauhiainen et al. 2019; http://urn.fi/urn:nbn:fi:lb-2019060602 (February 4, 2020).
49 Texts from the project “Corpus of Mesopotamian Anti-witchcraft Rituals” are included in our data but are not included in Korp. The reason for this is that they have been compiled from various different documents, and therefore the JSON files in Oracc do not contain the transliteration of the words (which is needed for Korp). At the same time, Korp contains some Akkadian texts that are not used in this article, most notably the lexical lists and bilingual texts.
metadata categories) of the word or words of interest. One can search for individual words or co-occurrences of two words within a certain distance from each other. After examining the contexts in Korp, certain possibilities became more probable than others. We then returned to the graphs to develop the idea further and used Korp at the same time to question and examine our analyses of contexts in graphs. (6) The final results of this analytical process were then reviewed in light of previous lexical research (see Section 2).

5 Data-Driven Results

PMI analyzes words that often appear close to our five verbs of interest. This is to say, it quantifies co-occurrence patterns. Therefore, these collocates reflect the vocabulary context in which our words of interest occur.

FastText is designed for handling large datasets efficiently, and instead of analyzing each word attestation in the corpus separately, it uses predictive models to build word vectors. It excels in building accurate word vectors in corpora of billions of words, but our tiny corpus of slightly more than a million words poses challenges for fastText. Relatively few words of synonymous or similar meaning appear on the lists of word vectors similar to the fear words under scrutiny. Instead we get many words which are similar to the PMI results, such as words that appear together (e.g., in formulaic expressions or lists).

Figures 18.1 and 18.2 give an example of what a simplified network of words looks like and exemplify the differences between the two methods. The actual
graphs we have produced in Gephi are too large to be easily reproduced in print. Nonetheless, even when comparing these simplified figures, one notices that because PMI requires the words of interest to appear close to each other, there are fewer connections between fear verbs than with fastText, which looks for words that appear in similar contexts as our words of interest.

As outlined above, our analysis is based on many different sources: fastText and PMI analysis is produced for two separate datasets (Level 1 and Level 2), and for each of these we have produced graphs for the 10 closest words and the 50 closest words. In the interests of clarity, we present below only our main observations. The intent here is to provide the reader with enough details to support our conclusions, although the main support for our arguments is provided by the PMI and fastText calculations and the graphs themselves.50

In the sections below, we discuss each of the five fear verbs in turn. Under each subheading we begin by discussing the verb and its derivatives (Level 1) and conclude by looking at the picture that emerges from Level 2 analysis, after the derivatives are combined into the verbal root. Each section is concluded by a reflective discussion that takes into account the earlier analyses of CAD (see Section 2).

5.1 palāḫu

To start with the Level 1 results, both PMI and fastText results offer a unanimous picture of the noun pulḫu. It denotes the terror caused by the Assyrian king and the god Aššur who accompanies him on the battlefield. Pulḫu is nearly always attested together with the word melemmu “radiance” (54 out of 58 occurrences) and almost as often with saḫāpu “to envelop, overwhelm” (48/58). A Korp search reveals that the words are attested in a common phrase in Neo-Assyrian royal inscriptions, of which a royal inscription of Sennacherib gives a good example: “Fear (pulḫu) of my lordly brilliance (melemmu) overwhelmed (saḫāpu) Lulī, the king of the city Sidon, and he fled afar into the midst of the sea and disappeared” (RINAP 3 23). PMI and fastText capture this relationship very well, as melemmu and saḫāpu have unusually high PMI scores in this dataset (0.042 and 0.028), and the cosine similarities show that the word vectors of melemmu (0.93) and saḫāpu (0.88) are very similar with that of pulḫu.

50 The graphs have been created with the Gephi software (downloaded from https://gephi.org). The full networks are available in our Zenodo repository (http://doi.org/10.5281/zenodo.3634325) in the “network-data” folder, whereas the visualizations of the networks as PDF files are in the “visualizations” folder.
The noun *puluḫtu* resembles *pulḫu*. They share *melemmu* and *saḫāpu* on their lists of the ten closest PMI collocates and most similar word vectors in fastText. Moreover, the word *namurratu* “radiance” also has a word vector which is similar to both *puluḫtu* (cosine similarity 0.77) and *pulḫu* (0.77). Co-occurrence numbers and PMI scores show, however, that unlike *pulḫu*, *puluḫtu* is not attested in a single phrase or set expression. There are several relatively frequent PMI collocates (e.g., *ḫātu* “terror” and *rašubbatu* “awesomeness”) which suggest that *puluḫtu* was widely used to denote fear and awe. A Korp query confirms this; *puluḫtu* is found in different text genres and contexts, denoting the fearsomeness of a god or king. The meaning appears to be close to *melemmu* and *namurratu*. Korp also reveals a context which is not apparent in the PMI and fastText results: in Neo-Babylonian royal inscriptions, *puluḫtu* means reverence for the gods (e.g., Schaudig 2001 no. 2.14).

The verb *palāḫu* itself appears in a number of contexts. It denotes a person’s anxiety and distress, reverence for gods, fear on the battlefield and performance of prebendary service. The lack of a single definite context is shown by the low PMI scores and very diverse lists of collocates and similar words in the PMI and fastText results.

There are four clearly distinct contexts for *palāḫu*. First, it is used to denote fear or anxiety in the context of personal distress, as suggested by the high-ranking PMI collocates *nakādu* “to beat, be anxious” and *naquttu* “distress.” According to Korp, they appear in contexts like “I was worried (*nakādu*), afraid (*palāḫu*), and hesitant about renovating that temple” (*RINAP* 4 57). Second, *palāḫu* has the meaning “to revere.” The verbs *sullû* “to appeal, pray” and *suppû* “to pray” have word vectors similar to *palāḫu*. Korp shows that they all appear in contexts depicting reverent behavior toward the gods (e.g., *RINAP* 5 Ashurbanipal 3; Schaudig 2001 no. 2.8). Third, *palāḫu* denotes fear experienced on the battlefield. The PMI collocates *namurratu* “radiance,” *kakku* “weapon” and *šitmuru* in the expression *tāhāziya šitmuri* “my stormy onslaught” regularly appear in the royal inscriptions of pre-Sargonid Assyrian kings (e.g., *RIMA* 3 A.0.102.2). Fourth, PMI collocates reveal a very specialized meaning of *palāḫu* in Hellenistic texts, “to perform prebendary service” (e.g., *OECT* 9 16). Here *palāḫu* is used together with the verb *rasānu* and its derivative noun *rāsinūtu*, both denoting the performance of prebendary service in this context.51

The adjective *palḫu* has two distinct meanings, “afraid” and “reverent,” according to our PMI results. The first meaning is characterized by two derivatives

---

51 See *CAD* R, 180–182.
of the fear verb adāru, šutadduru and adru. Both belong to the top three on the list of palḫu’s collocates, with šutadduru having a high PMI score (0.029). These three words appear to be nearly synonymous in contexts which characterize a person’s fear of evil signs and conditions (e.g., STT 2 231). The second meaning of palḫu, “reverent,” is suggested by the PMI collocates putuqqu “attentive” and wašru “submissive” in Neo-Babylonian royal inscriptions (Schaudig 2001 no. 2.8). In the fastText results, the most similar words to palḫu constitute a very heterogeneous group. Some of them relate to the meaning “afraid” (niqittu “anxiety,” but also “damage” in building accounts), others to the meaning “reverent” (e.g., kamāsu “to kneel”).

When the derivatives of palāḫu are combined with the verbal root on Level 2, vocabulary related to the fearsome radiance of the king and gods is most visible in the results. The words melemmu, namurratu and saḫāpu rank very high in both the PMI and fastText results. The contexts related to fearsome radiance on the battlefield, primarily attested in royal inscriptions, affect the fastText results in particular. The word vectors of the fear verb adāru and the anger verb ezēzu are similar to palāḫu, and these three words cluster together with other anger words and numerous battle-related words in the fastText graphs. At the same time, words relating to other semantic contexts of palāḫu and its derivatives are present in the PMI or fastText results, including rasānu and suppû, which relate to prebendary service and reverence of gods, respectively.

Our results are well in line with the picture that emerges from Akkadian dictionaries. According to the dictionaries, palāḫu seems to convey various aspects of fear and its use is not limited to a particular context. In our results, the frequent use of the verb and some of its derivatives in royal inscriptions emphasizes contexts related to battle. At the same time, two very different meanings of palāḫu, “to revere” and “to perform service,” are quite apparent in our results.

In many cases, palāḫu and its derivatives can be translated as “to fear” and “to be afraid,” but the meaning “to revere” is clearly present in our results as well. Conforming with the dictionaries, our results show that palāḫu and its derivatives pulḫu and puluḫtu denote fear of the king, gods and their numinous radiance (melemmu and namurratu), especially on the battlefield. Pulḫu and puluḫtu very often appear in these contexts, the former denoting the fear or terror caused by royal or divine radiance and the latter having a meaning close to melemmu and namurratu. At the same time, the verb palāḫu and the adjective palḫu are used to express personal anxiety and fear in different contexts. A rather different context for palāḫu and palḫu is that of reverence for the gods. Elements of fear were naturally involved in Mesopotamian divine-
human relationships, but the Akkadian usage of *palāḫu* seems to emphasize pious and respectful deeds and attitudes toward the gods. Finally, in Hellenistic texts, *palāḫu* refers to the performance of prebendary service.

Some smaller semantic nuances, such as respect toward one's parents, were not identified in our results. At the same time, connections to bodily locations, such as *libbu* ("heart"), which are evident in the dictionaries, are not visible in our results. This can be due to our corpus and its unbalanced distribution of genres (see Section 3) or the nature of our statistical method, which provides us with an aggregate perspective of the data. It is also possible that bodily locations of fear are far less often specified linguistically than we might expect and thus are not present in the results of the statistical analysis.

5.2 śaḫātu
There are two observations that can be made regarding the verb *šaḫātu* in Level 1. The first is confirmed on all four analytic layers. The verb *šaḫātu* is distinct from its derivative *šaḫtu* ("pious, reverent"), which clusters with *rāmu* ("to love") and the words derived from *rāmu*. The derivative *šaḫtu* appears to be used in a very specific context in our corpus. The majority of attestations derive from the king's titulary in royal inscriptions, surrounded by collocates such as *putuqqu* ("constantly attentive"), *wašru* ("submissive") and *narāmu* ("loved one"). These and several other related words (*rā'imu"one who loves," migru "favorite" of a god, *na'du"attentive/reverent," pāliḫu "who reveres/is respectful of" a god/king, *mutnennû "pious") describe the king's connection to deities, as can be seen from Korp. This observation strongly supports the translations of *šaḫtu* in all the major dictionaries, suggesting that the word denotes fear in the sense of being respectful toward the gods.52

One further observation can be made for Level 1 based on the PMI analysis. The collocates of the word *šaḫātu* itself are not very informative because there are three homonymous words attested, which in this level are all treated as one word. Nonetheless, the homonym *šaḫātu* "to be afraid" appears primarily in two contexts, which are repeated several times in the corpus. The collocates *anāqāte* ("she-camels"), *bakru* ("foal") and *ibilu* ("camel") relate to Tiglath-pileser's account of Samsi, the queen of the Arabs, who brought the king animals as a tribute because she was afraid. In several inscriptions of Esarhaddon, the king’s worry about rebuilding Babylon is a recurrent theme.

---

52 AHw, 1132: "demütig"; CAD Š/1, 101: "reverent, humble"; CDA, 347: "reverent, humble."
The collocates nakādu (“to be anxious”) and dikuggallu (“chief justice”) originate in this context.

In Level 2, where the derivatives are connected to the verbal root, the results are similar to Level 1. The connection of šaḫātu to ṛāmu (“to love”) is visible throughout, due to the strong connection of šahtu to ṛāmu in Level 1. There is thus little to add to the analysis of Level 1. The PMI analysis suggests some faint connections to palāhu and galātu, but these do not emerge as statistically significant connections. In terms of the fastText results, the cluster of ramû and šaḫātu appears in the neighborhood of anger verbs, but this is probably explained by the appearance of all these words in the royal inscriptions.

The overall result of our analysis is that the verb šaḫātu and its collocates (šahtu in particular) are tightly tied to the genre of royal inscriptions. This could also be seen in Korp when we examined the occurrences of the verb by genre. The clear majority of the occurrences (šaḫātu with the translation “be(come) afraid”) appear in royal inscriptions. In the dictionary analysis (see Section 2), a more varied picture emerges. Naturally, the connection between šaḥtu and ṛāmu would not be visible under the lemma šaḫātu in the dictionaries. Nonetheless, it seems that the difficulties we had with fastText and PMI for finding clear contexts for šaḫātu match the difficulties of pinpointing these contexts in the dictionaries. At least in our data it seems that this is a rarely used verb for expressing fear and worry in the specific genre of royal inscriptions.

5.3 adāru

Regarding Level 1, we made some observations about the verb adāru and its derivatives, ādiru and adru. The derivative noun adirtu, closely related to gilittu and pirittu, will be discussed in Section 5.4.

For adāru, the homonyms confuse the results of the analysis in Level 1, where homonyms have not been separated. The verb adāru A “to be disturbed, restless, obscured, worried” explains the words relating to astronomical phenomena (antallâ “eclipse” and collocates related to the points of the compass) in our results. The verb adāru B “to fear” relates to battle and campaign descriptions of Assyrian kings, such as the fear of snow (šalgu) or battle against the Assyrian king (the related collocates šuribatu “terror” and šuttinnu “bat”). In the N stem, the meaning of the verb shifts toward “being restless” or “raging” (the collocate huliam “helmet”). Furthermore, adāru has ties to šamru and ezzu (both “furious”) with further indicative words of martial contexts, such as tāḫāzu (“battle”) and tību (“attack”). These martial overtones are emphasized in the derivative form ādiru, which is closely connected to fear in battle, as an epithet of the Assyrian king (lā ādiru tuqumti “fearless in battle”).
The derivative *adru* has a rich homonym context,\(^5\) but the meaning of *adru* as “afraid” is reflected in the collocates *palḫu* (“frightened”) and *ṣutadduru* (“constantly terrified”). This group of three words originates from a number of incantation rituals in which a man says “I am frightened, I am afraid and I am constantly terrified.”\(^4\)

In Level 2, the process of separating homonyms and combining derivatives led to clear results of a martial context for *adāru*. Words with a martial theme connect it with two verbs related to anger: *ezēzu* (“be angry”) and *šamāru* (“to rage”) as well as (weakly) with *palāḫu* (“to fear”). Such words are, for example, *tāḫāzu* (“battle”), *tību* (“attack”), *tuquntu* (“battle”) and *ummānu* (“military force”). Furthermore, the terror caused by evil magic connects it to *galātu* and, to a lesser extent, *parādu* (e.g., the collocates *arurtu*, *šudlupu*, and *ḫuṣ ḫīp libbi* in texts related to anti-witchcraft rituals).

When we compare this with the results of CAD, we see that there, too, the homonyms of *adāru* are an issue. However, whether or not *adāru* A “to be disturbed, restless, obscured, worried” and *adāru* B “to fear, to respect” should be understood as the same verb is not relevant for this article, as we focus in any case on the meaning of “fear” of *adāru*. CAD suggests that commonly the verb has clear martial and battle connotations—this was very apparent in our results as well. Additionally, other causes for fear include natural elements like waves and snow, which did show up in the PMI analysis. However, there are two clear deviations from the image painted by the dictionaries. First, the context of terror caused by evil magic does not appear in dictionary analysis as a major context. Second, the dictionary meanings also list many fairly general meanings (and meanings similar to nuances of *palāḫu*): common objects are gods, oaths and parents, and the meaning can also have a positive connotation of respect regarding these. The verb in the dictionaries has many different translations as well, from intense fear to worry, sadness and annoyance. These meanings are not readily visible in our analysis.

These differences could be due to the composition of our data, but they could also be because our methods are better at picking up meanings that have a clearly defined lexical environment. Meanings and nuances which appear in many different text types and different contexts naturally have a varied set of words surrounding them. Therefore, such meanings are more difficult to pick up with statistical methods that analyze the lexical context in which the words appear. This becomes a challenge especially when the data (like the data that

---

53 In Korp, *adru* appears 183 times; the homonyms “threshing floor,” “dark,” and “eclipsed” account for a great majority of occurrences.

54 E.g., *palḫāku adrāku u šutaddurāku* sTt 2 23: rev. 9–10.
we have) includes text genres (such as royal inscriptions) that have a lot of
genre-specific vocabulary in them. This might explain why the contexts relating
to martial themes are prominent in our results.

5.4 galātu and parādu
In Level 1, both PMI and fastText provide a similar view of galātu and parādu. Although the verbs themselves appear to differ in meaning and contexts of usage, their derivatives gilittu and pirittu are regularly used together to denote fear or terror. The most common context in which gilittu and pirittu are attested together is the so-called disregard formula in oracular queries in which they appear together with the verb parādu. After derivatives are combined with roots, our results in Level 2 are dominated by the contexts in which gilittu, pirittu and parādu are attested together. Therefore, our discussion below focuses on Level 1.

The verb galātu has two basic meanings in our results, “to fear” and “to tremble.” This observation is well in line with the dictionaries consulted in Section 2. The verb is attested in several contexts, none of them frequently repeated. A number of the most similar words in fastText (e.g., ebēṭu “to swell up” and akālu in the sense of “to hurt”\(^{55}\)) suggest a context of divination and medical concerns. Galātu is indeed frequently attested in omens and incantation rituals, where it can denote fear caused by a snake (STT 2 321) or a rearing horse (TCL 6 9) but also the concrete trembling of a sick person (STT 1 91 +). Our PMI results suggest a number of additional contexts for galātu, including fear caused by the “dignity” (bāštu) of divine statues (RINAP 4 48). Galātu and palāḫu also occur together: in SAA 5 202, the former denotes the act of terrorizing and the latter refers to the fear caused by that. Finally, galātu can also denote the state of terror caused by the actions of people such as kings or military leaders (SAA 13 181) and criminals (SAA 19 188).

The verb galātu has a single frequently attested derivative, the noun gilittu. It clusters closely together with adirtu, parādu and pirittu in the PMI and fastText graphs. In particular, gilittu and pirittu are very strongly connected to each other; their PMI score (0.038) is one of the highest in our results, and there are no two word vectors more similar to each other (cosine similarity 0.93). These numbers suggest that gilittu and pirittu are very regularly used as a pair and that they are practically synonymous. Despite having weaker scores, adirtu and parādu are also closely associated with gilittu and pirittu.

---

\(^{55}\) Because of a mistake in the Oracc data, the normalization ikkalāšu was analyzed instead of the lemma akālu.
Gilittu and pirittu co-occur in two contexts. The tight connection between gilittu, pirittu and parādu is explained by the so-called disregard formula in oracular queries. An example of such a formula is: “Disregard that I, the haruspex your servant, am dressed in my ordinary soiled garments, have seen fear and terror (gilittu and pirittu) at night, or jumbled (parādu) the oracle query in my mouth. Let them be taken out and put aside!” (SAA 4 129, partially restored). It is noteworthy that parādu denotes unclear speech, not fear, in this context. Other relatively close PMI collocates of gilittu, pirittu and parādu occur in these contexts as well (for example, tamītu “oracle query” and aršu “dirty, soiled”).

Moreover, gilittu and pirittu are very similar to adirtu, a derivative noun of adāru. This is very clearly shown by the PMI and fastText results. The words are not attested together in the disregard formula, but adirtu occurs together with gilittu and/or pirittu in anti-witchcraft rituals in particular. In an anti-witchcraft ritual called Maqlû, these three words appear together with aššuštu “affliction,” arurtu “famine, trembling” and ḫūṣ ḫīp libbi “depression,” referring to a very distressed condition (Abusch 2015, tablet V). The words “terror” and “fear” seem to be legitimate English translations for adirtu, gilittu and pirittu.

The verb parādu is mainly attested in the disregard formula in our dataset, which is very clearly shown in the PMI and fastText results. A handful of other occurrences were studied using Korp. The word appears in medical omens as a symptom (either translated as “tremble” or “to be terrified”; e.g., SpTU 4 152), and it can denote fear experienced by mortals and gods (e.g., RINAP 4 106).

Regarding the derivative pardu, it is very frequently used in the expression “frightening dreams,” which explains the strong connection it has to šuttu (“dream”). The word has clear connotations related to magic and divination (e.g., ittu “sign”), and therefore it follows the general flavor of the verb parādu.

Since parādu and its derivatives have such a consistent context in Level 1, it is not surprising that the Level 2 results bring little new information. After derivatives are combined with roots, the contexts in which gilittu, pirittu and parādu are attested together overshadow the results.

To sum up, on all levels of analysis galātu and parādu cluster close to each other in a relatively well-defined connection. To a great extent this seems to be because of their connection via the pair gilittu and pirittu. The contexts suggest that these are close to being a compound expression. The verbs galātu and parādu themselves are similar as well, as both carry connotations of trembling in sickness, in addition to being afraid. Considering the genres where these two verbs and their derivatives appear (e.g., omens, incantations, healing literature), it seems plausible that these words were reserved for very specialized scribal use.
Based on dictionary analysis (see Section 2), we consider the root meaning of both verbs to be “negative movement” (e.g., “tremble”) and “disturbing the correct state of things.” These connotations are visible in our results, both in the concrete physical state in a medical sense and in the general context of rituals, healing and divination. The connection to medical conditions and divination is clear in the dictionary analysis as well.

6 Discussion

An important topic of this article has been the challenge of accessing the emotional lives of individuals in the past via their language. The semantic lexical fields of five Akkadian verbs that are commonly translated with “to fear” in English (adāru, galātu, palāḫu, parādu and šaḫātu) formed an entry point through which we have traced how the language users themselves would have connected these words to each other and to other words. Space here does not permit a full comparison to the semantic field of fear in the English language, but even a cursory glance reveals some differences. The clear positive sense of respect (palāḫu, adāru) is not represented by the English “fear” nor is the connotation of trembling and “fear” as a sickness or a symptom (particularly galātu, parādu, adāru). Additionally, the very specific idioms (e.g., “pay off a debt” for palāḫu) are naturally not visible in English “fear.” Finally, some of the verbs are used in a very limited context (e.g., šaḫātu, parādu) and therefore a proper English translation is very difficult to come by. This could be a matter of colloquial language vs. literary style. At the same time, the context of war (particularly palāḫu, adāru), as well as many of the less frequent contexts, can probably be found in an analysis of modern English as well. Furthermore, at least the Akkadian palāḫu can also be used to express less than dire circumstances (i.e., worry, just like “fear” in English).

The scope of the article has been limited out of necessity and this brief study cannot answer the many challenges of studying ancient emotions. As noted at the beginning of this article, the concept of “emotion” itself is not universal and the existence of universal emotional states is debated. Furthermore, there are many other ways of expressing the emotion of fear than these five verbs.

56 Although the English expression “fear of God” exists, it has clearly been transferred to English from Biblical Hebrew.
In this article, our modest and concrete goal has been instead to examine how emotions were expressed linguistically via these particular words and examine how the native Akkadian users themselves might have perceived the relationships between these five words and their relationships to other words. This aim rests on the idea that the meaning of any word lies in its relationship with other words. By using language technological methods (PMI and fastText), we traced semantically relevant properties of the five fear verbs with other words of Akkadian. This enabled us to create for the first time a semantic map of how these verbs relate to each other and other lexemes in our corpus. This provided us with tools to reflect on the semantic domains of the words beyond what is covered by existing research, even comprehensive volumes like the CAD.

In the article, we have discussed previous lexical research on “fear” in Akkadian (Section 2), described our dataset (Section 3), and outlined our research process and methods (Section 4) in detail. The results of the quantitative analysis are discussed in Section 5. In this final section, we summarize our results from the perspective of our main aims. The aims of our research have been twofold: 1) to develop methods that allow us to construct semantic maps of lexemes, thus getting closer to understanding the semantic domains in the mind of Akkadian language users, and 2) to use these methods to contribute something new to Assyriological research regarding lexical semantics of fear. We have contributed to these two goals through three distinct categories: 1) improving our methods, 2) examining “meaning” as lexeme networks, and 3) analyzing quantitative meaning and usage patterns of words.

6.1 Methods

As in all historical research, we are constrained by our source material. In our case, this is the text material that is available in Oracc. The existing genres influence the results strongly, as do the methodological choices that we have made.

Working on this article has shown us that in the future it will not be necessary to process the text data quite so much. Most of the interesting differences could be perceived immediately after the first round of processing, before combining the nouns with their verbal roots (see Section 3.2). The overall results highlight the difference between Level 1, with its basic simple modifications, and Level 2, where roots were combined with their derivatives and homonyms separated. Although both were useful, Level 1 should be our primary concern in future work, since a verb and its derivatives do not necessarily share the same semantic field. The process of taking homonyms into account was labor-intensive and does not seem to be strictly necessary in the future.
Still, this might not be true for some specific lexemes with commonly appearing homonyms (see adāru, for instance).

As far as the differences between PMI and fastText analysis are concerned, the variances proved to be interesting but not radically different. These methods offer two complementary images of our data, but probably because of the smallness of our data, fastText did not highlight very different lexemes when compared to PMI. However, Figures 18.1 and 18.2 above exemplify (in a simple way) that because PMI requires the words of interest to appear close to each other, there are fewer connections between fear verbs than with fastText, which looks for words that appear in similar contexts. These complementary views provide the best possible results.

Although we only selected verbs denoting fear that had a frequency of more than twenty instances in our dataset, during analysis we noticed that for words that appear less than fifty times in our material (e.g., galātu, piritu and šaḫātu), simply using the myriad search functions in Korp is a powerful research tool. Although the text corpus in Korp was not entirely identical with the text corpus used for this article (see Section 3), the results of Korp queries were often essential for our analysis. For example, the results of the Korp search demonstrated that the verb šaḫātu “to fear” is almost exclusively used to express fear in the genre of royal inscriptions. Certainly in Akkadian some genres are so specialized that they might constitute a context (a prototypical scenario where a word appears) in themselves. For example, we would classify the formulaic expressions in extispicy queries as such, but royal inscriptions are too varied to be considered a single context.

6.2 Meaning in Networks of Lexemes
The statistical results (PMI and fastText) on how specific words relate to each other give new perspectives on semantics and the relations between lexemes of fear in a way that has not been previously possible. Figures 18.1 and 18.2 above illustrate this to a small degree, although we again refer to the full dataset (available in the Zenodo repository). For a concrete example, the close connection between certain anger and fear verbs (especially in martial contexts) was very apparent in our results. The relationship of šaḫātu to love (through piousness toward gods) is worth mentioning here as well. Furthermore, parādu and galātu could be identified as semantically similar words.

See also Steinert’s contribution in this volume and her observation regarding fear-inspiring anger in warfare.
There are several interesting future research topics related to this approach. A potentially fruitful future topic could be to investigate, for example, how the lexicon of royal inscriptions differs from the lexicon of other text genres. In the results of Section 5, this genre figured prominently, and it could skew the results if we were interested in the “common” everyday use of the words. In other work currently in preparation, we intend to do more thorough analysis on the implications of genre regarding words relating to love.58

Another topic for more detailed research would be to study “bodily emotions.” Keywords on that topic, such as libbu “heart,” were not prominent in this article because of our focus on the five fear verbs. Interesting results could possibly be gained by selecting body parts as a topic of inquiry for our quantitative methods.

6.3 **Usage Patterns and Quantitative Meaning**

Although our statistical methods cannot significantly improve the existing lexical semantics (esp. CAD, but also other philological work on Akkadian lexical semantics) in terms of adding new nuances and new “meanings” to already known ones, what this approach can do is analyze the quantitative “meaning” of a word of interest.

We do not see it as a major drawback that not all dictionary meanings can be found with fastText or PMI. This would probably be the case even if we had the same text material at our disposal as the dictionaries had, as some of the meanings are attested so rarely. This relates to a larger question: is the meaning of a word the sum of all its potential meanings or is it defined by its most common usage? Arguments can be made for both, but a quantitative approach is better at answering the latter question. Therefore, instead of limiting ourselves to cataloguing possible translation meanings of a word, a quantitative approach can trace patterns of use for it. In this article we have been able to trace both common and rare patterns of usage.

Quantitative analysis is good for finding out the clearest context for a given word, that is, its most common pattern of use. The methods that we have applied seem to be especially good at identifying contexts that have specialized vocabulary and/or an easily recognizable setting (e.g., warfare). This explains at least partially why the royal inscriptions are so prominent in our results. Furthermore, some word pairs often appear together, and in very specific lexical contexts. Therefore, the methods easily recognize them. In this study, *gilittu*
and *pirittu* (seen in the expression “fear and terror”) form one such pair. At the same time, this tendency of the method suggests that the verbs for which it was difficult to find clear lexical contexts might have had more versatile uses in the lexicon than those with very clear contexts.

The results regarding *galātu* and *parādu* highlight the method’s ability to trace rare patterns of usage as well. It seems that these verbs expressing fear were known to and used only by a small minority of scribes. One of the strengths of statistical analysis is that one can thereby make fairly firm conclusions about the usage of words within the limits of the data.

Although dictionaries already document the meaning of a word well, with our method one can quantify these meanings and their contexts, tracing the common and uncommon patterns of use within Akkadian language. This suggests interesting possibilities for examining the sociolinguistic aspects of language, for example, where and when a certain vocabulary is used. We hope that our results will raise new questions and provoke readers to study some of these words from a new perspective with the help of the dataset we have published on Zenodo.

**Acknowledgements**

We gratefully acknowledge the funding received from the Academy of Finland that has made this article possible. Furthermore, we thank the Open Richly Annotated Cuneiform Corpus (Oracc) steering committee, in particular Niek Veldhuis, who provided access to Oracc data in JSON format. We are indebted to everyone who has been involved in making this research data available. This naturally includes the authors of the original publications, but also the people who have made the data Oracc-compatible and enriched it through lemmatizations and by adding other metadata. As the total number of people involved would amount to hundreds (the current number of individual subprojects in Oracc is approximately seventy), it would be impractical to list them all individually. However, much of our data is Neo-Assyrian, and much of that Neo-Assyrian linguistic data was created in the State Archives of Assyria project. We gratefully acknowledge SAA, created by Simo Parpola and his team and later developed into the State Archives of Assyria online (http://oracc.museum.upenn.edu/saao/) by Karen Radner and her team. We would also like to thank Ulrike Steinert for kindly sharing with us her draft article (appearing in this volume) and commenting on our draft. We are also grateful for feedback from Mikko Luukko, Niek Veldhuis and the anonymous reviewer, as well as Albion M. Butters for checking the English. Finally, we would like to thank the partic-
participants of the History of Emotions Symposium (Helsinki, November 7–8, 2018) for their feedback on an earlier version of this article.

**Abbreviations**

AHw von Soden, W. 1965, 1972 and 1981. *Akkadisches Handwörterbuch*. Wiesbaden: Harrassowitz.

CAD Oppenheim, A.L. and E. Reiner et al. (eds) 1956–2010, *The Assyrian Dictionary of the University of Chicago*. Chicago; Glückstadt: The Oriental Institute; J.J. Augustin.

CDA Black, J.A., A.R. George and J.N. Postgate (eds) 1999, 2002. *A Concise Dictionary of Akkadian* (Santag 5). Wiesbaden: Harrassowitz.

eSAD Supplement to the Akkadian Dictionaries. Leipzig.

FastText *Library for Efficient Text Classification and Representation Learning*. https://fasttext.cc.

JSON JavaScript Object Notation (open-standard file format).

OECT Oxford Editions of Cuneiform Texts. Oxford: Clarendon.

Oracc Open Richly Annotated Cuneiform Corpus.

PMI Pointwise Mutual Information.

RIMA The Royal Inscriptions of Mesopotamia, Assyrian Periods. Toronto: University of Toronto Press.

RINAP The Royal Inscriptions of the Neo-Assyrian Period. Winona Lake, IN: Eisenbrauns.

SAA State Archives of Assyria. Helsinki.

SAAo *State Archives of Assyria online*. http://oracc.museum.upenn.edu/saao/corpus.

SpTU Spätbabylonische Texte aus Uruk. Berlin: Mann.

STT Gurney, O.R., J.J. Finkelstein and P. Hulin. 1957, 1964. *The Sultantepe Tablets* 1–2. London: The British Institute of Archaeology at Ankara.

TCL Textes cunéiformes, Musée du Louvre. Paris: Geuthner.

**Bibliography**

Abusch, T. 2015. *The Magical Ceremony Maqlû: A Critical Edition*. Ancient Magic and Divination 10. Leiden: Brill.

Alstola, T., Zaia, S., Sahala, A., Jauhiainen, H., Svärd, S. and Lindén, K. 2019a. “Aššur and His Friends: A Statistical Analysis of Neo-Assyrian Texts.” *Journal of Cuneiform Studies* 71: 159–180.
Alstola, T., Zaia, S., Sahala, A., Jauhiainen, H., Svärd, S. and Lindén, K. 2019b. “Aššur and His Friends: A Statistical Analysis of Neo-Assyrian Texts” (Version 1.0) [Data set]. Zenodo. http://doi.org/10.5281/zenodo.2620131.

Alstola, T., Jauhiainen, H., Svärd, S., Sahala, A. and Lindén, K. in preparation. “Digital Approaches to Analyzing and Translating Emotion.”

Baroni, M., Dinu, G. and Kruszewski, G. 2014. “Don’t Count, Predict! A Systematic Comparison of Context-counting vs. Context-predicting Semantic Vectors.” In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics Baltimore, Maryland, USA, June 23–25, 2014, edited by K. Toutanova and H. Wei, 238–247. Stroudsburg: Association for Computational Linguistics. http://www.aclweb.org/anthology/P14-1023.

Bastian, M., Heymann, S. and Jacomy, M. 2009. “Gephi: An Open Source Software for Exploring and Manipulating Networks.” International AAAI Conference on Web and Social Media. https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154.

Blois, R. de. 2008. “Semantic Domains for Biblical Greek: Louw and Nida’s Framework Evaluated from a Cognitive Perspective.” In Foundations for Syriac Lexicography III: Colloquia of the International Syriac Language Project, edited by J. Dyk and W. van Peursen, 265–278. Perspectives on Syriac Linguistics 4. Piscataway, NJ: Gorgias Press.

Blondel, V.D., Guillaume, J.-L., Lambiotte, R. and Lefebvre, E. 2008. “Fast Unfolding of Communities in Large Networks.” Journal of Statistical Mechanics: Theory and Experiment: P10008.

Bojanowski, R., Grave, E., Joulin, A. and Mikolov, T. 2017. “Enriching Word Vectors with Subword Information.” Transactions of the Association for Computational Linguistics 5: 135–146. https://arxiv.org/pdf/1607.04606.pdf.

Church, K.W. and Hanks, P. 1990. “Word Association Norms, Mutual Information, and Lexicography.” Computational Linguistics 16: 22–29.

Daille, B. 1994. Approche mixte pour l’extraction automatique de terminologie: Statistiques lexicales et filtres linguistiques. PhD diss., Université Paris 7.

Eustace, N., Lean, E., Livingston, J., Plamper, J., Reddy, W.M. and Rosenwein, B.H. 2012. “AHR Conversation: The Historical Study of Emotions.” The American Historical Review 117(5): 1487–1531. https://doi.org/10.1093/ahr/117.5.1487.

Firth, J.R. 1957. “A Synopsis of Linguistic Theory, 1930–1955.” In Studies in Linguistic Analysis, edited by J.R. Firth, 1–32. Oxford: Blackwell.

Gruber, M.I. 1990. “Fear, Anxiety and Reverence in Akkadian, Biblical Hebrew and Other North-West Semitic Languages.” Vetus Testamentum 40(4): 411–422.

Jacomy, M., Venturini, T., Heymann, S. and Bastian, M. 2014. “ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software.” PLoS ONE 9: e98679.

Jaques, M. 2006. Le vocabulaire des sentiments dans les textes sumériens: Recherche sur le
lexique sumérien et akkadien. Alter Orient und Altes Testament 332. Münster: Ugarit-Verlag.

Jauhiainen, H. and Alstola, T. forthcoming. “Fast(Text) Analysis of Mesopotamian Divine Names.” In CyberResearch on the Ancient Near East and Neighboring Regions (vol. 2). Evaluating new tools and methods for Archaeological Data, Objects, Texts, and Digital Archiving, edited by V. Bigot Juloux, S. Matskevich, and A. di Ludovico. Leiden: Brill.

Jauhiainen, H., Sahala, A. and Alstola, T. “Open Richly Annotated Cuneiform Corpus, Korp Version, May 2019” [text corpus]. Kielipankki. http://urn.fi/urn:nbn:fi:lb-2019060601.

Manning, C.D. and Schütze, H. 1999. Foundations of Statistical Natural Language Processing. Cambridge, MA: MIT Press.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. and Dean, J. 2013a. “Distributed Representations of Words and Phrases and their Compositionality.” In Advances in Neural Information Processing Systems 26, edited by C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani and K.Q. Weinberger. https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf.

Mikolov, T., Chen, K., Corrado, G. and Dean, J. 2013b. “Efficient Estimation of Word Representations in Vector Space.” Last modified September 7, 2013. https://arxiv.org/abs/1301.3781 (accessed April 16, 2019).

Mirguet, F. 2016. “What is an ‘Emotion’ in the Hebrew Bible? An Experience that Exceeds Most Contemporary Concepts.” Biblical Interpretation 24(4–5): 442–465.

Noack, A. 2009. “Modularity Clustering Is Force-directed Layout.” Physical Review E 79: 026102.

Pennington, J., Socher, R. and Manning, C.D. 2014. “GloVe: Global Vectors for Word Representation.” In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), edited by A. Moschitti, B. Pang, and W. Daelemans, 1532–1543. Association for Computational Linguistics.

Plamper, J. 2015. The History of Emotions: An Introduction. Oxford: Oxford University Press.

Rendu Loisel, A.-C. 2010. Bruit et émotion dans la littérature akkadienne. PhD diss., University of Geneva.

Riley, A.J. 2017. Divine and Human Hate in the Ancient Near East: A Lexical and Contextual Analysis. Perspectives on Hebrew Scriptures and its Contexts 25. Piscataway, NJ: Gorgias.

Role, F. and Nadif, M. 2011. “Handling the Impact of Low Frequency Events on Co-occurrence Based Measures of Word Similarity: A Case Study of Pointwise Mutual Information.” In Proc. International Conference on Knowledge Discovery and Information Retrieval (KDIR-2011), edited by J. Filipe and A. Fred, 226–231. https://doi.org/10.5220/0003655102260231.
Schaudig, H. 2001. *Die Inschriften Nabonids von Babylon und Kyros’ des Großen samt den in ihrem Umfeld entstandenen Tendenzschriften: Textausgabe und Grammatik*. Alter Orient und Altes Testament 256. Münster: Ugarit-Verlag.

Svärd, S., Jauhiainen, H., Sahala, A. and Lindén, K. 2018. “Semantic Domains in Akkadian Texts.” In *CyberResearch on the Ancient Near East and Neighboring Regions*, edited by V. Bigot Juloux, A.R. Gansell and A. di Ludovico, 224–256. Digital Biblical Studies 2. Leiden: Brill.

Wasserman, N. 2016 *Akkadian Love Literature of the Third and Second Millennium BCE*. Leipziger Altorientalistische Studien 4. Wiesbaden: Harrassowitz.