Using Word Embeddings for Identifying Emotions Relating to the Body in a Neo-Assyrian Corpus

Ellie Bennett and **Aleksi Sahala**
University of Helsinki
eleanor.bennett@helsinki.fi
aleksi.sahala@helsinki.fi

Abstract

Research into emotions is a developing field within Assyriology, and NLP tools for Akkadian texts offers new perspectives on the data. We use PMI-based word embeddings to explore the relationship between parts of the body and emotions. Using data downloaded from Oracc, we ask which parts of the body were semantically linked to emotions. We do this through examining which of the top 10 results for a body part could be used to express emotions. After identifying two words for the body that have the most emotion words in their results list (*libbu* and *kabattu*), we then examine whether those emotion words were indeed used in this manner in the Neo-Assyrian textual corpus. The results indicate that of the two body parts, *kabattu* was semantically linked to happiness and joy, and had a secondary emotional field of anger.

1 Introduction

The study of emotions in ancient Near Eastern cultures has grown in the past five years, with two edited volumes covering the topic (Hsu and Llop Raduà, 2021; Sonik and Steinert, 2022). A key question has been how ancient languages used the body to express emotions - or the 'embodiment of emotions'. For Akkadian, this question has been addressed by Ulrike Steinert (Steinert, 2022, 2021). Her work involved the traditional approach of close reading Akkadian texts and close analysis of the most important Akkadian dictionaries to identify words relating to emotions, and how the body was used to express them. She identified 22 Akkadian words that refer to parts of the body that were used to express emotions (Steinert, 2022). As Steinert's results were reflective of Akkadian language as a whole, we were interested in whether her results could be replicated for a textual corpus from a more narrow timespan - namely the Neo-Assyrian period (c. 934-612 BCE).[1]

Neo-Assyria is the best represented time period currently available on the Open Richly Annotated Cuneiform Corpus (Oracc). We can therefore use word embeddings to gather quantitative data of whether the body words identified by Steinert to express emotions were used similarly to emotion words during the Neo-Assyrian period. In addition, word embeddings will allow us to determine whether any of these body words can be considered as part of a semantic field of emotions.

2 Data

The data for this project consists of two sections: the corpus of Neo-Assyrian texts, and the list of Akkadian words that relate to parts of the body.

2.1 Neo-Assyrian corpus

The textual corpus was downloaded from Oracc.[2] The texts were selected according to metadata tags found in the Oracc data. We chose a 'fuzzy' approach to the data, and cast a wide net in order to include as many Neo-Assyrian texts written in Akkadian as possible. We therefore included texts with the following tags for language and time period:

- 'Akkadian'; 'akkadian'; 'Akkadian with Sumerian incipits'; 'Akkadian, Aramaic?'; 'Akkadian, with Aramaic epigraph'; 'Akkadian?'; 'Assyrian'; 'Akkadian, Aramaic'
- 'Neo-Assyrian'; '9th/8th century'; '8th/7th century'; '9th century'; '7th century'; '8th century'

The data is in a word per line format, where every word is represented in the following lemmatised

[1] There is growing research into emotions in Neo-Assyrian material (Valk, 2016; Battini, 2022; Bach, 2022; Nadali, 2022; Schaudig, 2022; Bonatz, 2022; Morello, 2022), but scholarship is still limited with regards to embodied emotions.

[2] This was done through a remix of the following script, which resulted in a lemmatised version of the Oracc dataset in line with Oracc standards: https://github.com/niekveldhuis/compass.

Proceedings of the Ancient Language Processing Workshop associated with RANLP-2023, pages 193–202,
held in Varna Bulgaria, Sept 8, 2023.
https://doi.org/10.26615/978-954-452-087-8_022

form:[3]

lemma[guideword]EPOS[4]

Stop words are given as '<stop>',[5] unlemmatised words as '_',[6] and '#' to indicate the end of a text.

The data underwent a process of minimal cleaning to remove duplicates that appeared due to spelling errors, resulting in a corpus of 7,969 texts, 1,014,890 tokens, and 19,436 unique word forms.[7]

2.2 Word selection

We referred to Ulrike Steinert's work on embodied emotions in Akkadian to select the words that were most likely to be similar to emotion words in the Neo-Assyrian dataset (Steinert, 2021, 2022). We selected 22 words relating to the head (5 words), torso (or the whole body) (3 words), organs (6 words), and limbs (8 words).[8] They can be viewed in Table 1.

3 Word Embeddings

Word embeddings represent words as real-valued vectors in a multi-dimensional vector space. In simple terms, this vector space can be understood as a matrix, where the rows represent each word in the corpus and the columns represent abstractions of their co-occurrences with other words. This property makes word embeddings useful for measuring lexical similarity: if two words A and B are similar to each other in meaning, they likely occur in similar contexts, and thus their vectors should show higher similarity to each other. The most closely related words (often called *nearest neighbors*) can be computed from the vectors by using cosine similarity (Equation 1).

$$\cos(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} \quad (1)$$

In word embedding calculations, this metric has a theoretical upper bound of 1 in case the words are perfect synonyms, and a lower bound of 0 in case the semantic relationship between the words is independent.[9]

Some of the better known tools for building word embeddings are Word2vec (Mikolov et al., 2013) and fastText (Bojanowski et al., 2017), which both use shallow neural networks to learn the word embedding space from the corpus. While the neural network based methods are powerful in large datasets, these methods are known to have issues in smaller datasets comprising only a million words (Jungmaier et al., 2020). For low-resource languages and small corpora, word embeddings can be successfully built by using count-based methods, such as word association measures combined with matrix factorization (Levy et al., 2015; Jungmaier et al., 2020). The count-based method used in this paper produces vectors only for words that are documented in the data set. Therefore the similarity between out-of-vocabulary words cannot be measured like in fastText.

PMI-embeddings is a Python script for building count-based word embeddings for low-resource languages.[10] It combines several features from various research papers on count-based word embeddings that utilize Pointwise Mutual Information (PMI) (Church and Hanks, 1990) and Singular Value Decomposition (SVD). Briefly described, PMI-embeddings builds a sparse matrix of the vocabulary that encodes statistical significance of co-occurrences in terms of PMI within a symmetric context window of arbitrary size. This sparse matrix is then truncated into desired dimensionality (typically between 60 and 300) using SVD, yielding the final vector space that can be saved into the standard Word2vec format to be used in various NLP applications.

The script allows tuning several hyperparameters to find optimal settings for the given dataset. These include, but are not limited to, Dirichlet Smoothing

[3] As Akkadian is a highly inflected language, and the research was not focused on syntax or morphology, we worked from the lemmatised version of the dataset. This follows previous Assyriological research projects based on Oracc data (Svärd et al., 2020; Sahala and Svärd, 2021; Bennett, 2023)

[4] This is how the Akkadian words will appear in the tables of this contribution for the ease of non-Assyriological readers, but in the main body we will be using Assyriological standards of italicising the lemma.

[5] A full list of these can be found in the accompanying Zenodo repository, the url of which can be found in Appendix A.

[6] An unlemmatised word in Oracc can be due to many factors, such as a broken text (indicated by '*x*' in transliterations), or a word with uncertain meanings.

[7] The dataset is part of the supplementary material, which is described in Appendix A. The corpus can be broken down according to genre in the following manner: letters (33%), royal inscriptions (17%), transactions (15%), scholarly texts (14%), and the remaining 21% was a mix of administrative, religious, legal, political, literaturary, and untagged texts.

[8] We did not include bodily emissions in this research, but could be an interesting future field of emotions research (Sonik, 2022a).

[9] Negative scores up to -1 are also possible depending on how the word embeddings are created, but in this paper discussion on negative values is not necessary.

[10] https://github.com/asahala/pmi-embeddings/

Head	Body	Organs	Limbs
īnu[eye]N	pagru[body]N	kabattu[liver]N	ahu[arm]N
pānu[front]N	šīru[flesh]N	kalītu[kidney]N	durā'u[arm]N
pû[mouth]N	zumru[body]N	karšu[stomach]N	idu[arm]N
qaqqadu[head]N		libbu[interior]N	izīru[arm]N
rēšu[head]N		qerbu[centre]N	kirimmu[(crook-of)-arm]N
		ṣurru[interior]N	kittabru[arm]N
			purīdu[leg]N
			zāqu[arm]N

Table 1: The 22 Akkadian words relating to the body selected for this study organised according to whether they refer to the head, the whole body or torso, the organs, or limbs.

(Turney and Pantel, 2010; Jungmaier et al., 2020), several variants of PMI such as shifted and context distribution smoothed PMI (Levy et al., 2015) with various different shifting mechanisms, Dynamic Context Window that gives less significance for co-occurrences that are further apart (Sahlgren, 2006), and Context Similarity Weighting (Sahala and Lindén, 2020) that downsamples noise caused by duplication and repetitiveness in the dataset, which is a problem especially in formulaic Mespotamian royal inscriptions.

If a human-evaluated gold standard for semantic similarity is available, well performing parameters can be searched by using the *hypertune.py* script distributed with PMI-embeddings. Currently the default settings of PMI-embeddings have been defined for the first millennium Akkadian texts using a work-in-progress gold standard based on independent word similarity rankings done by five Assyriologists. This gold standard is distributed as a part of PMI-embeddings.[11]

3.1 Parameters

The default parameters use a vector dimensionality of 300 and the following features for calculating the sparse PMI matrix:

- Shifted PMI with a shift value of 7 using the formula implemented in Jungmaier et al. 2020. This allows some co-occurrences to exist in the sparse matrix even if they are not statistically significant. Typically all statistically independent co-occurrences would be disregarded, which may be harmful in sparse datasets.

- Symmetric dynamic context window of three words, meaning that the co-occurrences are calculated within a span of three preceding and following words to the center word, and that the co-occurrence frequencies are reciprocals of their distance to the center word giving less importance to words that co-occur farther away from each other.

- Context Similarity Weighting with a k-value of 3, meaning that co-occurrences in repetitive or partially repetitive contexts are downsampled with a weight risen to the power of 3. This gives significant penalty to co-occurrences also in partially repetitive contexts.

4 Results

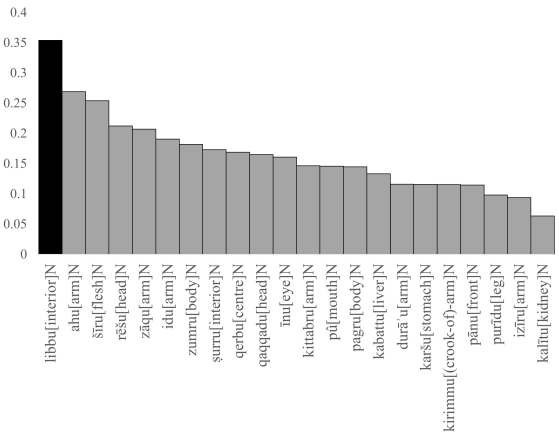

Figure 1: Ranges of the differences between the first and last cosine score in the top ten results for Akkadian body words. *Libbu* is highlighted in black.

We queried the resulting .vec file for each of the 22 Akkadian words relating to the body for the 10 words whose vectors were the most similar to

[11] https://github.com/asahala/pmi-embeddings/tree/main/eval built in cooperation with the University of Helsinki, the LMU Munich and the University of California, Berkeley.

the Akkadian body word, as well as their cosine similarity score.

Libbu had the largest range of cosine similarity scores between the first and tenth result out of all the results lists. *Kalītu* ("kidney") had the smallest difference in cosine score in its top ten results. The mean of the differences in cosine similarity score of the top 10 results was 0.16, and the median was 0.15, indicating most of the body words queried had a similar range of usage in the Neo-Assyrian corpus.

The difference between the body words with the highest (*libbu*) and second highest (*ahu*) range in cosine scores was 0.085 - the largest difference between the cosine ranges (Fig. 1).[12] The large range in cosine score in comparison to the other results lists suggests that *libbu* was used in many different contexts, whereas the smaller difference between *kalītu*'s first and tenth results suggest it had a more specific usage in Neo-Assyrian texts.

The words in the results lists for each body word were then compared to a master Excel file that classified emotions according to 18 different emotion categories: sadness, distress, suffering, anger, happiness, schadenfreude, pleasure, fear, hate, love, desire, disgust, sympathy, envy, pride, surprise, shame, sexual arousal.[13]

If a word in the results lists occurred in this Excel file, the principle emotional field was imported to the results list for analysis. This highlighted the words in the results list that could be used to denote emotions. This was our first indicator of whether a body word was part of an emotional semantic field.

4.1 Words relating to the organs

The group of words referring to organs (column 3 in Table 1) was most likely to have a word in its results list that could express emotions. Of the word relating to organs, *libbu* ("interior") and *kabattu* ("liver") had the highest number of emotion words in their results list (5 and 4 respectively, as seen in Tables 2 and 3).[14]

[12]Removing *libbu* from the results does not change the mean and median of the ranges of cosine scores for the top 10 results of each body word.

[13]The full list of emotional fields can be found in the accompanying Zenodo repository (Appendix A).

[14]Although the guideword translates this as "interior", *libbu* is actually a difficult word to translate. The precise definition is a general sense of "inside", including "inside" a house or the body. The lack of precision for precisely *which* organ inside the human body *libbu* referred to has led to scholars translating the term variously as "heart", "liver", and "mind" (CAD L: 164 *libbu*). We decided to keep the original sense of

libbu		
Word	**Score**	**Emotion**
kindu[earth]N	0.786	None
hīpu[break]N	0.525	Sadness
gilittu[terror]N	0.496	Fear
qīlu[burning]N	0.462	None
haluppu[(a-tree)]N	0.458	None
hūṣu[pain]N	0.457	Sadness
hibṣu[swelling]N	0.453	None
terku[blow]N	0.443	Fear
zenûtu[anger]N	0.441	Anger
papānu[(a-kind-of-rush)]N	0.432	None

Table 2: The ten words most similar to *libbu* in the Neo-Assyrian corpus.

kabattu		
Word	**Score**	**Emotion**
teṣû[defecate]V	0.793	None
ṣabru[blinker]AJ	0.788	None
alālu[sing-a-joyful-song]V	0.770	Happiness
elēṣu[swell]V	0.715	Happiness
mussahhiru[turned-to-someone)]AJ	0.701	None
hadû[be(come)-joyful]V	0.697	Happiness
mušnēšu[keeping-alive]AJ	0.683	None
aggu[furious]AJ	0.661	Anger
qurruru[(meaning-unknown)]AJ	0.661	None
epēru[feed]V	0.660	None

Table 3: The ten words most similar to *kabattu* in the Neo-Assyrian corpus.

The results for *libbu* have a surprising feature. The scores between the first and second results (kindu and *hīpu*, respectively) have the biggest difference of any two scores in the results lists. Kindu, a Sumerian term found in bilingual texts twice, has a score of 0.786, and *hīpu* has a score of 0.525 - a difference of 0.261. As kindu only occurs twice in the corpus, and *libbu* occurs 5,710, this result is either a quirk of the *PMI embeddings* script, or - more likely - is representative of the wide usage range of *libbu* in Neo-Assyrian texts.

The emotion words in the results list of *libbu* are overwhelmingly negative: *hīpu* ("break") and *hūṣu* ("pain") were used in the expression of sadness (Wende, 2022).[15] *Gilittu* ("terror") and *terku* ("blow") were used to express fear, and *zenûtu* ("anger") was used for anger (Wende, 2022; Bach, 2022; Svärd et al., 2020).[16]

In comparison, the emotion words similar to *kabattu* were mostly positive (Table 3). *Alālu* ("sing a joyful song"), *elēṣu* ("swell"), and *hadû* ("be joyful"), were all used to denote happiness (Wende, 2022; Bach, 2022).[17] The only outlier is *aggu* ("furious"), which was used to express anger (Wende, 2022; Bach, 2022).[18]

Of the other words relating to organs, *qerbu* ("centre") had the third highest number of words relating to emotions in its results list, with 2 results.[19] *Zā'eru* ("hostile") was the 4th highest result, with a cosine similarity score of 0.508.[20] It was used in the expression of despisement or hate. The ninth word in the results was *maqtu* ("fallen one"), with a cosine similarity score of 0.471.[21] The principal emotional field of this word is surprise.

Finally, *ṣurru*'s ("interior") results list also featured a word relating to emotions: *rūbu* ("anger"), ranked ninth, with a cosine similarity score of 0.669.[22]

4.2 Words relating to the limbs

Of the eight words relating to the limbs, four had words that could be used to express emotions in their results lists.

Purīdu ("leg") had the highest number of emotion words in its list, with *muštarhu* ("presumptuous one") ranked third (0.543 cosine similarity) and *petû* ("open") ranked fourth (0.519 similarity).[23] *Muštarhu* expresses pride whereas *petû* expressed happiness.

Two of the other words for parts of the arm had positive emotions appear in their results list. *Zāqu* ("arm") saw *ṭābu* ("good") as the most similar word, with a cosine similarity score of 0.751, and was used to express happiness.[24] The ninth result for *kirimmu* ("crook of the arm") was *narāmtu* ("beloved", cosine value 0.508), used to express love.[25]

Lastly, the seventh result for *ahu* ("arm") was *adirtu* ("gloominess") (cosine similarity of 0.546), and was used to express sadness.[26] This was the only result connected to a negative emotion for the words relating to limbs.

4.3 Words relating to the head, torso, and whole body

None of the Akkadian words relating to the head, torso, or the whole body had a word denoting emotions in their results lists.

5 Analysis

The only body words that had more than three results which could be used to express emotions were *libbu* and *kabattu*. Previous scholarship regarding Neo-Assyrian embodied emotions has identified these were often used in phrases to express emotions (Bach, 2022; Luukko, 2021; Morello, 2022; Sonik, 2022b). The results appeared to therefore align with current Assyriological research.

Thus far, we have identified words in results lists that *could* be used to express emotions. Many of these words can be translated in different ways and have alternative usages. This section will assess whether the words in the results lists for *libbu* and *kabattu* were indeed used to convey emotions in the Neo-Assyrian corpus. This will indicate whether

the word, and have simply left it as the guideword "interior". CAD K: 11 *kabattu*.

[15] CAD H: 195 *hīpu*; CAD H: 260 *hūṣu*.

[16] CAD G: 71 *gilittu*.

[17] CAD A1: 331 *alālu*; CAD E: 88 *elēṣu*; CAD H: 25 *hadû*.

[18] CAD A1: 150 *aggu*.

[19] CAD Q: 216 *qerbu*.

[20] CAD Z: 14 *zā'iru*.

[21] CAD M1: 254 *maqtu*.

[22] As with *libbu*, *ṣurru* was also used to denote a general "inside" of the human body, and cannot be connected to any individual (or group of) organs as we understand them today. CAD Ṣ:259 *ṣurru*; CAD R: 400 *rūbu*.

[23] CAD P: 517 *purīdu*; CAD M2: 286 *muštarhu*; CAD P: 338 *petû*.

[24] CAD Z: 64 *zāqu*; CAD Ṭ: 19 *ṭābu*.

[25] CAD K: 406 *kirimmu*; CAD N1: 342 *narāmtu*.

[26] CAD A1: 205 *ahu*; CAD A1: 127 *adirtu*.

emotions in general were part of the semantic field for either *libbu* or *kabattu*. It will also examine whether the emotions conveyed with these words align with current Assyriological research into embodied emotions.

5.1 Emotion words similar to *libbu*

Libbu is consistently identified in Assyriological scholarship as the main seat of emotion, and is described in Akkadian phrases as where anger, fear, and joy were felt (Bach, 2022; Luukko, 2021; Steinert, 2022; Schaudig, 2022).

The first word relating to emotions in *libbu*'s results list was *hīpu* ("break"), which was the second overall result. In the Neo-Assyrian corpus, *hīpu* was attested 72 times, and was used in three ways: to describe a break in a text the scribe was copying;[27] to describe quarrying lapis lazuli;[28] the breakage of physical objects;[29] and to describe the breaking of the *libbu* as an expression of sadness. Only 3 attestations were to express emotion and were only found in the royal inscriptions of Assurbanipal in the construction *hīp libbu* ("broken *libbu*", often translated as "broken heart") (Wende, 2022).[30] Therefore, the primary usage of *hīpu* in the Neo-Assyrian corpus was not to express emotion, but when it was, it was done so only with *libbu*.

The second word on the results list that could express emotion was *gilittu* ("terror"). *Gilittu* occurs most frequently in the Neo-Assyrian corpus in disregard formulae of oracular queries (22 out of 27 attestations) (Svärd et al., 2020; Wende, 2022).[31] In all of the attestations, *gilittu* was used to express fear, and more specifically acute terror. As *gilittu* has a clear semantic field of fear in the Neo-Assyrian corpus, its appearance as third on the results list for *libbu* suggests that of the emotions, *libbu* was connected strongly to fear.

The third emotion word on *libbu*'s results list was *hūṣu* ("pain"). It is ranked sixth in the similarity rankings, and it is unsurprising to see *hūṣu* alongside *hīpu* in the results list for *libbu*, as these words were part of the compound expression *hūṣ hipi libbi* ("an emotional or physical pain within the *libbu*").[32] However, in the Neo-Assyrian corpus this phrase was not used to express an emotion, but in scholarly commentaries to describe an abdominal pain.[33] Thus, in the Neo-Assyrian corpus *hūṣu* connected *libbu* with a medical semantic field.

The fourth word relating to emotions was *terku* ("blow"). *Terku* could be used to describe a throbbing emotional response, but of the 14 Neo-Assyrian attestations this was only found in 1 text. The text was a letter to the Assyrian king and includes a description of someone dying from a throbbing *libbu* due to hearing the speech of the king.[34] The rest of the attestations were in two royal inscriptions and the omen series *Šumma tirku*. In all of these cases, *terku* was used to describe a physical dark spot on either human skin or lambs' wool.[35] Overall, even though *terku* could be used to express a strong emotion or physical response, in the Neo-Assyrian corpus it was used for anatomical descriptions.

The final word used to express emotions in the results list for *libbu* is *zenûtu* ("anger"). It had a clear connection with the emotional field of anger, but was only used in two royal inscriptions from the reign of Esarhaddon in passages explaining the anger of the god Marduk prior to destroying Babylon.[36] It followed the phrase *ēziz libbašu* ("his *libbu* was furious"), which explains it's appearance in the results list, and indicates *libbu*'s connection

[27] In order to aid non-Assyriologists, we refer the reader to the electronic versions of the texts hosted on Oracc. When possible, the urls also link to the exact location of the word under discussion in the text. For examples of *hīpu* used to decribe a break in an original tablet the scribe copied, this can be seen in the first three lines of http://oracc.org/blms/P394721, lines rev. ii 8-19 in http://oracc.org/cams/gkab/P338598, and four lines in http://oracc.org/dcclt/nineveh/P386432.

[28] For example http://oracc.org/rinap/rinap1/Q003421.4.5; http://oracc.org/rinap/rinap4/Q003232.125.4; and http://oracc.org/rinap/rinap4/Q003230.308.5.

[29] For example in a recipe for making glass in line rev. 37 in http://oracc.org/glass/P394484.

[30] http://oracc.org/rinap/rinap5/Q003819.5.2; http://oracc.org/rinap/rinap5/Q007602.249.2; http://oracc.org/rinap/rinap5/Q003710.875.3.

[31] For example, in http://oracc.org/saao/saa04/P238980.29.12, http://oracc.org/saao/saa04/P239001.16.1, and http://oracc.org/saao/saa04/P238965.22.1.

[32] CAD H: 260 *hūṣu*.

[33] For example, http://oracc.org/ccpo/P296515.8.1, http://oracc.org/ccpo/Q005179.6.1, and http://oracc.org/ccpo/P461217.7.2.

[34] http://oracc.org/saao/saa10/P313436.19.2

[35] For example, http://oracc.org/cams/gkab/P363488.12.2, http://oracc.org/rinap/rinap4/Q003388.5.5, http://oracc.org/rinap/rinap4/Q003387.7.7.

[36] http://oracc.org/rinap/rinap4/Q003335.18.1, http://oracc.org/rinap/rinap4/Q003342.9.7.

to the semantic field of anger.

Overall, only two of the words that appeared on the results list for *libbu* were used exclusively to express emotions in the Neo-Assyrian corpus: *gilittu* and *zenûtu*. They expressed fear and anger, respectively.

Traditional Assyriological research states *libbu* was the seat of joy based on phrases that locate joy within the *libbu* (Steinert, 2022; Luukko, 2021). Our results diverge from this perspective, and offer an alternative perspective of the emotions associated with *libbu*. Neither *gilittu* nor *zenûtu* were used to express joy, and none of the words in *libbu*'s results list that could be used to express emotions were used to express joy. Interestingly, 27 attestations pair *elēṣu* with *libbu*. *Elēṣu*'s absence on the results list for *libbu* further strengthens the argument that whilst joy might have been connected with *libbu*, *libbu* was not primarily associated with joy or happiness.

Moreover, whilst *libbu* was used in a similar manner to *gilittu* and *zenûtu*, the varied usage of other words in the result list suggests that emotions were not the principle semantic field for *libbu*. This was made even more clear once the results list was expanded to 20 results, as the results pointed to a semantic field more akin to illness and agriculture. These can be found in the supplementary material (Appendix A). These results from close reading also corroborate the suggestion from the ranges in cosine similarity scores (Fig. 1), which suggested *libbu* would have a broader semantic field than other Akkadian body words in this study.

Emotions were therefore not the principle semantic field for *libbu*, and the results list speak to it's diverse usage in the Neo-Assyrian corpus. However, when it was used to express emotions, it was used to express fear and anger.

5.2 Emotion words similar to *kabattu*

In Assyriological scholarship, *kabattu* is consistently identified as an important part of the body where emotions were felt, especially for feeling happiness and joy (Bach, 2022; Sonik, 2022b).

Of the emotion words in the results list for *kabattu*, *alālu* ("to sing a joyful song") was the most similar, and was ranked third in the results list. Of the seven attestations in the Neo-Assyrian corpus, it was used to express the gods' joy five times,[37] and the joy of the king twice.[38] *Alālu* was found in literary texts, royal inscriptions, and a hymn. All of these genres drew upon literary topoi, which suggests *alālu* was imbued with metaphorical meaning (Wende, 2022).

Elēṣu ("to swell")was the second word related to emotions in the results list for *kabattu*, and was ranked fourth. It is mostly attested in Neo-Assyrian royal inscriptions (35 attestations out of 37 in the corpus), and was largely used to describe the happiness and good moods of the gods and the kings (Bach, 2022).[39] In six texts it was paired with *kabattu* (such as *ēteliṣ kabattī*), desribing how joy made the liver swell.[40]

Hadû ("to become joyful") was the third word related to emotions in the results list for *kabattu*, and was ranked sixth. *Elēṣu* and *hadû* were two of four words identified by Mikko Luukko that expressed joy or happiness in Assyrian archival material (Luukko, 2021). *Hadû* specifically related to divine joy and the joy of the king, which aligns with Luukko's view that the happiness and joy of superiors was a concern for subordinates (Luukko, 2021).[41] There is also an interesting usage where *hadû* was used to describe the happiness of fathers in literary texts.[42] This points to a particular type of joy that *hadû* expressed, and *kabattu* could be semantically connected to not just joy, but a specific kind of joy.

The final word relating to emotions that was on the results list for *kabattu* was *aggu* ("furious"),

[37] http://oracc.org/saao/saa03/P334929.85.1, http://oracc.org/rinap/rinap2/Q006494.149.2, http://oracc.museum.org/cams/gkab/P338326.37.2, http://oracc.org/cams/gkab/P363581.33.4, and http://oracc.org/cams/gkab/P338328.30.4.

[38] http://oracc.org/rinap/rinap2/Q006488.194.7, http://oracc.org/rinap/rinap2/Q006483.336.1.

[39] For example http://oracc.org/rinap/rinap2/Q006488.168.9, http://oracc.org/rinap/rinap4/Q003230.496.4, and http://oracc.org/rinap/rinap5/Q007625.37.4.

[40] http://oracc.org/cams/anzu/Q002771.60.5, http://oracc.org/ribo/babylon6/Q006325.77.2, http://oracc.org/rinap/rinap2/Q006596.62.4, http://oracc.org/rinap/rinap2/Q006482.325.2, http://oracc.org/rinap/rinap2/Q006483.189.8, and http://oracc.org/rinap/rinap2/Q006605.77.1.

[41] For example, http://oracc.org/riao/Q004661.10.1, http://oracc.org/saao/saa09/P337163.23.7, http://oracc.org/saao/saa01/P224485.22.1, and http://oracc.org/saao/saa04/P237053.13.8.

[42] http://oracc.org/atae/huzirina/P338675.56.2, http://oracc.org/cams/gkab/P338321.83.3, and http://oracc.org/cams/gkab/P338675.56.2.

which was ranked eighth. It was mostly used in royal inscriptions to describe how the wrath of the gods was not appeased.[43] Even though this was a different emotional field, it ties in to the theme of emotions of gods seen in the usage of the other emotion words similar to *kabattu*.

The findings align with the usage of *kabattu* in similar types of texts. Of the 176 attestations, 120 were in royal inscriptions, which were heavily inspired by literary topoi in order to express emotions as agreed upon by the king and the most senior circle of scribes (Bach, 2022).[44] Most of the attestations of *alālu*, *elēṣu*, *hadû*, and *aggu* were in similarly literary texts. In addition, the results demonstrate an overwhelming concern regarding the happiness of gods and kings, aligning with previous scholarship suggesting happiness was a key aim of Mesopotamian kingship (Morello, 2022; Luukko, 2021).

Overall, the results for *kabattu* included words that were used in texts like royal inscriptions which made extensive use of literary motifs. Of the words that were similar to *kabattu*, and could be used to express emotions, all were indeed used as such in the dataset. Three out of four were used to express happiness or joy. Therefore, *kabattu* was principally part of the semantic field of happiness and joy. The fourth result (*aggu*) suggests a secondary semantic field for *kabattu* was anger, based on these words' usage in royal inscriptions and literary texts.

6 Conclusions

PMI-embeddings has therefore proven to be an important tool to not only identify which areas of the body were most associated with emotions, but to identify which body parts were semantically connected to specific emotional fields in Neo-Assyrian texts.

We highlighted two words that were most similar in usage to words that could be used to express emotions: *libbu* and *kabattu*. Our much more limited list of Akkadian words for body parts that were semantically linked with emotions than Steinert's list suggests the Neo-Assyrian corpus has temporally-specific methods of embodying emotions. Future research could compare our results to datasets built from texts of a different period, such as Middle Assyrian. *Libbu* and *kabattu* were the two Akkadian words relating to the body that were most likely to be part of the semantic field of emotions, aligning with the findings of previous Assyriological research (Sonik, 2022b; Steinert, 2021; Luukko, 2021; Wende, 2022).

We have been able to corroborate results with close readings of texts for *kabattu*, and solidify that in the Neo-Assyrian texts a primary semantic field of the liver was not just emotions generally, but more specifically the semantic fields of happiness and anger.

We have also demonstrated how the results from *PMI-embeddings* can complement and add to close reading approaches in order to provide a more nuanced reading of Neo-Assyrian embodied emotions. We demonstrated that the emotion words most similar to the usage of *libbu* were in the emotional field of fear and anger, which does not align with traditional close reading approaches. Furthermore, our results demonstrate that emotions were only one facet of the many usages of *libbu* in Neo-Assyrian texts.

PMI-embeddings is therefore a tool that can both corroborate results from close-readings, but more importantly has provided new perspectives on how Neo-Assyrian texts embodied emotions. On the basis of this single case study, *PMI-embeddings* will become vital in the suite of tools for digital Assyriology.

Acknowledgements

This research was conducted as part of the project Embodied Emotions: Ancient Mesopotamia and Today, funded by the Finnish Cultural Foundation. The *PMI-embeddings* script was developed with the assistance of the Centre of Excellence Ancient Near Eastern Empires, funded by the Academy of Finland (decision number 352747).

We would like to thank Dr. Ulrike Steinert, who graciously allowed us access to the research of the project "Akkadische und Hethitische Emotionsbegriffe im Kontext" (AHEC).

This research could not have been carried out without the richly annotated dataset on Oracc, almost all of which was manually inputted. We would like to thank the Steering Committee of

[43] This was in eight out of 12 attestations. For example, http://oracc.org/rinap/rinap5/Q003705.538.3, http://oracc.org/rinap/rinap5/Q003703.316.3, and http://oracc.org/rinap/rinap5/Q003778.23.3.

[44] For example, http://oracc.museum.upenn.edu/rinap/rinap2/Q006596.62.4, http://oracc.museum.upenn.edu/rinap/rinap4/Q003286.275.1, and http://oracc.museum.upenn.edu/rinap/rinap5/Q007615.16.6.

Oracc, Jamie Novotny, Eleanor Robson, Steve Tinney, and Niek Veldhuis. We would also like to thank the PIs of the many Oracc sub-projects, and the scores of researchers who were involved in the lemmatisation and digitisation process.

Abbreviations

CAD: The Assyrian Dictionary of the University of Chicago

References

Johannes Bach. 2022. Emotions and Assyrian Kingship. In Karen Sonik and Ulrike Steinert, editors, *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 329–358. Routledge, London.

Laura Battini. 2022. Sennacherib's Sieges and Deportations Reliefs: How to Increase Emotions. *Avar: An Interdisciplinary Journal of Life and Society in the Ancient Near East*, 1(2):313–359.

Ellie Bennett. 2023. Age and Masculinities during the Neo-Assyrian Period. *Journal of Cuneiform Studies*, 75:123–154. Publisher: University of Chicago press.

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information.

Dominik Bonatz. 2022. Fear and Terror in Assyrian Palace Reliefs. In Karen Sonik and Ulrike Steinert, editors, *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 525–543. Routledge, London.

Kenneth Ward Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1):22–29.

Shih-Wei Hsu and Jaume Llop Raduà, editors. 2021. *The Expression of Emotions in Ancient Egypt and Mesopotamia*. BRILL.

Jakob Jungmaier, Nora Kassner, and Benjamin Roth. 2020. Dirichlet-smoothed word embeddings for low-resource settings. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 3560–3565, Marseille, France. European Language Resources Association.

Omer Levy, Yoav Goldberg, and Ido Dagan. 2015. Improving distributional similarity with lessons learned from word embeddings. *Transactions of the Association for Computational Linguistics*, 3:211–225.

Mikko Luukko. 2021. Expressions of Joy and Happiness in Neo-Assyrian. In Shih-Wei Hsu and Jaume Llop Raduà, editors, *The Expression of Emotions in Ancient Egypt and Mesopotamia*, pages 255–282. BRILL.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space.

Nathan Morello. 2022. Joy and Happiness in Mesopotamian Royal Inscriptions. In Karen Sonik and Ulrike Steinert, editors, *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 455–471. Routledge, London.

Davide Nadali. 2022. Shaming the Enemy in Assyrian Palace Reliefs and Royal Inscriptions. In Karen Sonik and Ulrike Steinert, editors, *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 614–627. Routledge, London.

Aleksi Sahala and Krister Lindén. 2020. Improving word association measures in repetitive corpora with context similarity weighting. In *Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2020, Volume 1: KDIR, Budapest, Hungary, November 2-4, 2020*. SCITEPRESS Science And Technology Publications.

Aleksi Sahala and Saana Svärd. 2021. Language technology approach to "seeing" in Akkadian. In *The Routledge Handbook of the Senses in the Ancient Near East*, 1 edition, pages 561–575. Routledge, London.

Magnus Sahlgren. 2006. The word-space model: using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces.

Hanspeter Schaudig. 2022. Anger and Hatred in Neo-Assyrian and Neo-Babylonian Royal Inscriptions. In Karen Sonik and Ulrike Steinert, editors, *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 631–647. Routledge, London.

Karen Sonik. 2022a. Emotions and Body Language. The Expression of Emotions in Visual Art. In Karen Sonik and Ulrike Steinert, editors, *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 269–325. Routledge, London.

Karen Sonik. 2022b. Emotions and Body Language. The Expression of Emotions in Visual Art. In Karen Sonik and Ulrike Steinert, editors, *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 269–325. Routledge, London.

Karen Sonik and Ulrike Steinert, editors. 2022. *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition. Routledge, London.

Ulrike Steinert. 2021. Pounding Hearts and Burning Livers: The Sentimental Body in Mesopotamian Medicine and Literature. In Shih-Wei Hsu and Jaume Llop Raduà, editors, *The Expression of Emotions in Ancient Egypt and Mesopotamia*, pages 410–469. BRILL, Leiden.

Ulrike Steinert. 2022. Emotion and the Body: Embodiment, Conceptual Metaphor, and Linguistic Encoding of Emotions in Akkadian. In *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 51–87. Routledge, London.

Saana Svärd, Tero Alstola, Heidi Jauhiainen, Aleksi Sahala, and Krister Lindén. 2020. Fear in Akkadian Texts: New Digital Perspectives on Lexical Semantics. In S-W Hsu and J. Llop-Raduà, editors, *The Expression of Emotions in Ancient Egypt and Mesopotamia*, number 116 in Culture and History of the Ancient Near East, pages 470–502. Brill, Leiden.

Peter D Turney and Patrick Pantel. 2010. From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*, 37:141–188.

Jonathan Valk. 2016. "They Enjoy Syrup and Ghee at Tables of Silver and Gold": Infant Loss in Ancient Mesopotamia. *Journal of the Economic and Social History of the Orient*, 59(5):695–749.

Janine Wende. 2022. Akkadian Emotion Terms. In Karen Sonik and Ulrike Steinert, editors, *The Routledge Handbook of Emotions in the Ancient Near East*, 1 edition, pages 150–170. Routledge, London.

A Supplementary Data

Supplementary data is available to view for this research. Accompanying files can be found via the following url:

`https://doi.org/10.5281/zenodo.8272738`

The data includes:

- A ReadMe file.
- The Neo-Assyrian corpus downloaded from Oracc.
- A .txt file listing the terms for the body.
- The Excel spreadsheet with the emotion words and their emotional categories.
- The vector files generated by the method outlined in the main text.
- An Excel file with the top 10 results for each word, their cosine similarity score, the frequencies for the words, and their principal emotional fields.
- All images used in this publication.