Meaning Change in Face Emojis in German Social Media Data

Geraldine Baumann and Tatjana Scheffler
Digital Forensic Linguistics
Department for German Language and Literature
Ruhr University Bochum, Germany
firstname.lastname@rub.de

Abstract

We present the first large-scale study of semantic change in face emojis on German social media data. We calculate and evaluate yearly word embeddings based on eight years of online communication (2015–2022), which we compare to analyze semantic shifts in 107 face emojis. We identify several distinct trajectories of semantic development for different emojis and conduct detailed case studies for four representative emojis illustrating them. Finally, we show that some emojis exhibit meaning change reflecting the *Apparent Time Hypothesis*, demonstrating age-based linguistic variation. Our findings provide new insights into the real-time evolution of emoji meanings and help delineate the limitations of distributional semantic representations for emoji (or word) meaning.

1 Introduction

Due to the absence of facial expressions and other non-verbal cues, emojis are frequently used to clarify messages and convey emotional nuances in digital communication (Boutet et al., 2023; Kralj Novak et al., 2015). Although the text remains identical, messages such as “It’s okay ☺️” and “It’s okay 😢” can communicate very different meanings. This highlights the important role emojis play in everyday communication and demonstrates the need to understand and model their interpretation.

Unlike words, the introduction date of emojis is tracked, and the vast amounts of available online data allow us to trace meaning shifts directly (Robertson et al., 2021). Recent studies have thus started adapting methods from semantic change detection to emojis, particularly for English-language data (Robertson et al., 2021; Arviv and Tsur, 2022). Computational methods such as time-specific embedding spaces have proven particularly useful. Meaning shifts between two vector spaces can, for example, be detected using the method described by Gonen et al. (2020), where the overlap between the lists of closest neighbors of a word in two embedding spaces is examined. However, no such studies on meaning shifts in face emojis have so far been conducted on German-language data.

Furthermore, it still remains unclear which emojis are more likely to undergo semantic change. A possible approach to this issue is offered by the *Apparent Time Hypothesis*, which argues that synchronous data can give information on diachronic change through differences between age groups (Bailey, 2004). While several studies have examined differences in age-based interpretation or usage of emojis (e.g., Teh, 2025; Boutet et al., 2024; Herring and Dainas, 2020), many emojis do not show any age-based variation, and no link has been established to meaning change to date.

The current study investigates semantic change in 107 face emojis using a corpus spanning from 2015 to 2022. Our research questions are as follows: 1) Which face emojis can be shown to have changed their meaning in German-language data? 2) Which patterns of meaning change throughout time-spans can we identify? 3) Do these shifts in meaning correspond to differences in interpretation of face emojis between different age-groups?

Our findings extend previous research based on English-language data and provide new insights into the semantic development of emojis. Furthermore, they contribute to a broader understanding of how social and demographic factors may influence semantic change in digital communication. Finally, we also show that distributional semantic representations as used here have limitations when analyzing certain expressive emojis such as ☺️. To aid future research, we make our scripts and results publicly available via an online repository[1].

Our main contributions are as follows:

- We adapt the method proposed by Gonen et al.

[1] Find our repository here: https://osf.io/st3h2/.

(2020) for diachronic analysis of emoji meanings in German-language data.

- We identify four different types of semantic trajectories for emojis over time and conduct four case studies on exemplary emojis demonstrating these.
- We examine the relationship between semantic change and age-based interpretation of emojis (= the *Apparent Time Hypothesis*).

2 Related Work

Many studies have investigated semantic change and developed methods to detect it. Among the most widely used are alignment-based approaches that compare embedding spaces from different time periods (Hamilton et al., 2016). These methods align the vector spaces using an orthogonal transformation that preserves similarities between vectors as much as possible. An alternative is to avoid explicit alignment altogether. For example, Gonen et al. (2020) compare a word's representations by measuring the overlap between its k nearest neighbors in independently trained embedding spaces. This approach avoids modifying the representations while producing results that are easily interpretable.[2]

While methods for semantic change detection exist, only a few studies have investigated meaning shifts in emojis, especially outside the context of English data. Teh (2025) examined whether differences between age groups' interpretations of emojis mirror linguistic shifts, expecting older participants to prefer old meanings, while younger participants have already adopted the newer meaning of an emoji. Participants were divided into age groups and were successively shown five emojis with the task to either choose from a given list of definitions or provide a free-form definition if none of the provided definitions was suitable. The participants were first shown the emojis by themselves and later in a fictional context based on the presumed use in youth language. Across all five examined emojis, older participants generally provided more literal and established meanings, while younger ones often gave new meanings. However, when provided with a context, all participants gave the intended, newer meaning, independent of age.

Robertson et al. (2021) trained month-based embeddings from 2012 to 2018 to investigate changes in the meaning of emojis. Using the *local neighborhood measure* allowed them to identify temporal shifts. They found, for example, a shift for the emoji *Basketball* (🏀) during the *NBL*-season in the United States, before which the emoji was more often used as a general representation of a ball. Despite finding emojis that changed their meaning during specific seasons, most emojis showed a stable representation across their examination period.

Month-based embeddings were also used by Arviv and Tsur (2022), spanning from May 2016 to April 2019. In addition, they also divided their training data according to the platform used (web interface, iOS, Android), training three embedding spaces for each year. They compared matrices of pairwise word and emoji distances and demonstrated that platform-dependent renderings can influence the interpretation and usage of emojis. For example, re-rendering the former *Pistol* emoji as a *Water Pistol* (🔫) caused people to use it more frequently in the context of video games instead of actual weapons.

While previous studies have demonstrated meaning shifts in emojis using English data, no comparable research has been conducted on German data. Furthermore, little is known about the properties that might make certain emojis more susceptible to semantic change. Age-based variation may provide a first approach to answering this question.

3 Data

Social media data is particularly suitable for emoji research due to the high frequency of emoji usage. Scheffler (2014) describes a method by which German-language posts from *Twitter* were collected between 2014 and the shutdown of the APIs in early 2023, resulting in a corpus containing more than 90% of German-language tweets from this period. Complete yearly datasets are available from 2015 to 2022, which, therefore, constitutes the study period of this paper. This data is the basis for the present study and is examined to address research questions (1), (2), and (3).

Since face emojis are the most frequently used and show distinct linguistic behavior (Grosz et al., 2023), we concentrate on only face emojis here. Scheffler and Nenchev (2024) conducted a norming study for all 107 face emojis available at the time, collecting affective, semantic, frequency, and descriptive norms from German speakers. The present work examines the same set of emojis. Since par-

[2] More information in Section 4.

ticipants also provided information about their age, this data will also be used to address research question (3). Emoji implementation dates were obtained from the *Unicode Consortium*[3] to account for emojis that were unavailable in earlier years.

Table 1 summarizes the corpus size and the number of tweets containing at least one of the 107 examined face emojis. Of these emojis, 73 were already available in 2015, the first year of our study period, the others were introduced during the time of collection.

Year	Tweets	with Emoji	%
2015	191,844,902	11,423,714	5.95
2016	218,698,289	14,498,616	6.63
2017	135,387,282	10,869,288	8.03
2018	152,737,929	13,485,274	8.83
2019	179,392,534	18,572,997	10.35
2020	203,491,368	22,109,541	10.87
2021	262,939,603	29,470,285	11.21
2022	274,890,001	39,421,157	14.34

Table 1: Corpus overview including the total number of tweets per year and the number of tweets containing a face emoji (with percentages).

4 Method

To detect semantic change across time, this study adopts the nearest-neighbor approach by Gonen et al. (2020). Unlike alignment-based methods (e.g., Hamilton et al., 2016), this approach does not require projecting embedding spaces into a shared vector space. Instead, semantic change can easily be estimated by comparing the k nearest neighbors of a word across two embedding spaces. In addition, since this approach not only provides similarity values between emojis-at-a-certain-time but also word lists of their overlapping neighbors, results are easily interpretable. The method assumes that low overlap between neighbor sets indicates a semantic difference.

For each word (w), the semantic change *score* is calculated as the negative size of the intersection between the sets of k nearest neighbors (NN) in two embedding spaces:

$$score^k(w) = -|NN_1^k(w) \cap NN_2^k(w)|$$

Using this method has the advantage that the meaning change score of a word is determined only by its own neighbors in the vector spaces, while the alignment of two vector spaces must account for all words in the vocabulary. For our diachronic analysis, semantic change was calculated between two embedding spaces based on data from consecutive years in the examination period.

Preprocessing was kept minimal to preserve as much of the original data as possible. Tweets were split into sentences at punctuation marks using regular expressions to avoid context shifts at sentence boundaries. The resulting sentences were tokenized, and mentions, hashtags, and links were removed to reduce noise and anonymize the data.

5 Training

To create the vector spaces, we used the *Skip-Gram* version of the *FastText* algorithm (Bojanowski et al., 2017) provided by *Gensim*.[4] Separate embedding spaces were trained for each year from 2015 to 2022. The window size was set to six, as this was shown to be among the best performing by Barbieri et al. (2016), and negative sampling was set to 10. The models were trained for 30 epochs. Based on Barbieri et al. (2016), emojis were treated as tokens during the training process.

6 Evaluation

We conduct a thorough quantitative and qualitative evaluation of the embeddings in order to ensure that they are good representations of emoji meaning. The quantitative test aimed to verify the quality of the embedding spaces as a whole, while the qualitative test was carried out to gain insight into the quality of the emoji embeddings specifically.

For quantitative evaluation, we used the German translation of the *Google semantic/syntactic analogy*-dataset (Köper et al., 2015), which provides 18,552 analogy tasks divided into 13 semantic and syntactic categories. Analogies were solved using the most similar function from *Gensim* and counted as correct if the expected answer was among the top 10 most similar. Table 3 (in Appendix A) presents the resulting accuracy scores divided into categories for each year.

All embedding spaces received an overall accuracy score of 0.63 or higher (Table 3 in Appendix A). Problems can be seen particularly with the categories *currency* and *city-in-state*. The latter category includes exclusively US-American examples

[3]Unicode 2025: "Emoji Versions, v17.0." URL: https://www.unicode.org/emoji/charts/emoji-versions.html

[4]https://radimrehurek.com/gensim/models/fasttext.html

that are arguably rarely discussed in German Twitter data. The former category was also shown to be among the worst-performing for the models by Köper et al. (2015) themselves. This could indicate that currencies and countries are not commonly discussed in similar contexts. The models performed particularly well on the tasks *capital-common-countries* and *plural-verbs*.

For the qualitative evaluation, we created yearly visualizations of the emoji embeddings, projected onto two dimensions using *t-SNE*, similarly to Barbieri et al. (2016). Figure 1 shows one example based on the year 2015. Several clusters of similar emojis can be seen, showing that related emojis were given similar representations. For example, in the top right corner, a cluster of emojis with kissing mouths and hearts can be seen, and to the left of it, a cluster of emojis with wide, open smiling mouths. Most emojis with downturned, sad mouths or tears can be found at the bottom towards the middle of the visualization. The other visualizations can be found in Appendix A (Figures 8–14).

Additionally, to support these findings, we created a second set of *t-SNE*-based 2D visualizations, where each emoji is represented using its valence rating from the norming study conducted by Scheffler and Nenchev (2024) (Figure 15 in Appendix A). Within each year, the visualizations show a consistent gradient from positive to negative emojis, showing that emojis with similar emotional valence (positive or negative) are located close to each other in the embedding spaces. We conclude that the trained embedding spaces are good representations of the language in general and of the emoji meanings in particular.

7 Diachronic Examination

Following the approach by Gonen et al. (2020), we examine the semantic change for each face emoji between two consecutive years in the examination period. For each emoji, we extracted the 1000 nearest neighbors from the embedding space and calculated the overlap between the neighbor sets of two years. The value of $k = 1000$ was adopted from Gonen et al. (2020) and chosen because our embedding spaces are based on a large dataset. Using larger neighborhoods allows for the detection of more subtle semantic shifts than smaller neighborhoods restricted to only the closest neighbors. Table 6 (in Appendix D) shows the overlap between all year pairs available for each emoji. A shortened overview sorted by years can be found in Table 2. A higher value (closer to 0) signifies little overlap between the closest neighbors in the two year-based vector spaces, indicating a potential meaning shift or unstable representation. The emoji *Smiling Face* (☺️) can, for example, always be seen among the emojis with the lowest overlap between year pairs (right side of Table 2). The highest overlap for this emoji can be seen between the years 2015 and 2016, with only 7 identical tokens in the top 1000 closest neighbors.

In contrast, a lower value (closer to -1000) shows a larger overlap and a lower probability for a meaning change between a year pair. An example is *Face with Tears of Joy* (😂); this emoji is consistently among the emojis with the most overlap between years (left side of Table 2). The highest overlap for this emoji can also be found in the year pair 2015/2016, with 669 out of the 1000 closest neighbors being identical. For this year pair, *Face with Tears of Joy* is also the emoji with the highest overlap of all examined face emojis.

8 Results

We can observe distinct patterns in the overlap values over time. Many of the emojis show a fairly stable representation and have high overlaps in the neighbor sets between each year pair. *Face with Tears of Joy* (😂), for example, is among the emojis with the highest overlap for all year pairs. In contrast, *Smiling Face* (☺️) is consistently among the emojis with the least overlap. This is surprising, since both emojis are among the ten clearest and most familiar face emojis in German (Scheffler and Nenchev, 2024). This would generally suggest stable and widely shared interpretations for both emojis.

The emoji *Face with Medical Mask* (😷) has a fairly stable representation but experiences a sudden decrease in overlap between 2019 and 2020. The emoji *Skull* (💀) exhibits reduced overlap values across multiple consecutive year pairs between 2016 and 2019. While the indication for a semantic shift for *Face with Medical Mask* is short-lived, the shift of *Skull* appears more gradual and sustained across time.

Figure 2 visualizes the overlap-score development for these four emojis[5]. The visualization shows a sharp temporary decrease for *Face with*

[5]A more detailed examination of these emojis follows in Section 10.

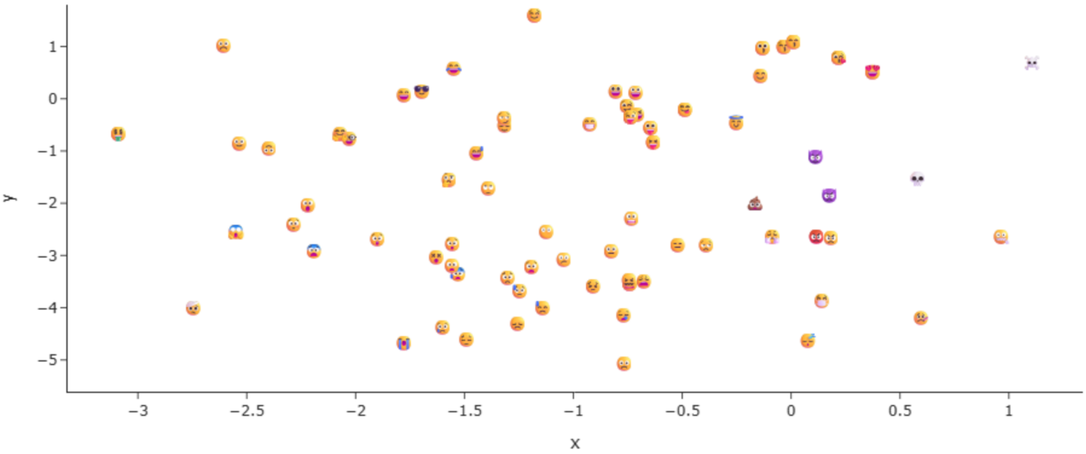


Figure 1: T-SNE visualization of face emoji embeddings (n=73) reduced to two dimensions based on the model for the year 2015.

Years	10 emojis with the most overlap	10 emojis with the least overlap
2015–2016	{😂: -669, 😅: -616, 😴: -608, 😃: -606, 😉: -603, 😣: -582, 😀: -578, 😆: -575, 😎: -570, 😜: -566}	{😔: -418, 🤑: -415, 🙄: -402, 🤔: -397, 😄: -379, 😍: -358, 😭: -348, ☹️: -12, ☠️: -8, ☺️: -7}
2016–2017	{😂: -581, 😴: -568, 🤒: -558, 😎: -537, 😅: -535, 😭: -527, 🤠: -527, 😜: -516, 😕: -501, 😣: -496}	{😏: -359, 👿: -337, 🤤: -264, 🤔: -261, 💀: -193, 😃: -190, 🤥: -172, ☹️: -8, ☺️: -4, ☠️: -3}
2017–2018	{🤣: -641, 😴: -627, 😂: -622, 😄: -598, 😎: -591, 🤕: -584, 🧐: -583, 🤒: -581, 🤗: -565, 🙂: -560}	{🤬: -379, 😔: -367, 🤔: -365, 🤪: -349, 😘: -295, 😃: -279, 💀: -265, ☠️: -65, ☹️: -10, ☺️: -1}
2018–2019	{🤣: -669, 😂: -647, 🤒: -609, 😴: -606, 🤓: -602, 😩: -599, 😫: -597, 😎: -583, 🤗: -563, 😆: -562}	{😈: -350, 🤠: -344, 😘: -318, 🥴: -305, 🥺: -301, 🤡: -139, 💀: -131, ☠️: -6, ☹️: -4, ☺️: -1}
2019–2020	{🤓: -667, 🥶: -626, 🤣: -612, 😴: -603, 😎: -593, 🤢: -591, 😫: -590, 😩: -587, 🤮: -577, 😂: -576}	{😙: -291, 😭: -280, 🥱: -265, 😗: -258, 😐: -242, 🤧: -203, 😷: -123, ☹️: -7, ☠️: -3, ☺️: 0}
2020–2021	{🤓: -675, 🤣: -654, 😴: -651, 😎: -633, 😫: -631, 🤢: -625, 😣: -622, 🤫: -617, 😂: -616, 🥴: -616}	{😓: -413, 😍: -413, 😮: -409, 🤥: -406, 🤔: -382, 😶‍🌫️: -68, ☠️: -12, 😵‍💫: -8, ☺️: -5, ☹️: -1}
2021–2022	{🤓: -695, 🤣: -672, 😂: -660, 💀: -648, 🥶: -625, 😴: -625, 😩: -625, 🤫: -623, 😫: -620, 😱: -615}	{😵‍💫: -7, ☺️: -1, 🥹: 0, 🫢: 0, ☠️: 0, 🫤: 0, 🫠: 0, 🫥: 0, 🫣: 0, 🫡: 0 }

Table 2: Ten emojis with the highest (left) and lowest (right) overlaps for each year pair.

Medical Mask, whereas *Skull* forms a prolonged plateau of lower overlap values. *Face with Tears of Joy* and *Smiling Face* represent comparatively stable high- and low-overlap patterns, respectively. Because these four emojis illustrate distinct developmental patterns (*stable representation*, *unstable representation*, *rapid semantic shift*, and *gradual semantic shift*), they are used as case studies in the following analysis.

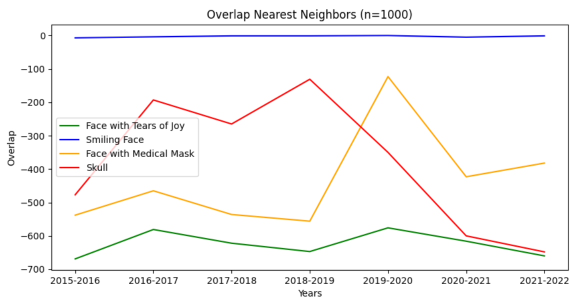


Figure 2: Visualization of the meaning overlap scores between the years 2015 and 2022 for four example emojis.

The following chapters aim to examine possible age-related reasons for meaning shifts in emojis, asking whether a relationship between semantic change and age-based interpretation of emojis can be found (Section 9). Furthermore, we provide detailed examples for the different trajectories shown in this section (Section 10).

9 Age-Based Meaning Shifts

While the method provided by Gonen et al. (2020) allows us to identify emojis that potentially undergo a meaning shift, it does not provide information on the reasons for a meaning shift. Under the *Appar-*

ent Time Hypothesis, linguistic differences between two age groups may indicate a change in progress, as older people are seen to exhibit an earlier language variant than younger people (Bailey, 2004). To address research question (3), we apply this hypothesis to the case of short-term emoji meaning change as observed in our data, hypothesizing that apparent time meaning differences in emojis correspond to real-time semantic change of the emojis. Such a connection could provide first insights into which emojis change their meaning. To investigate this possibility, we reanalyze the data from the norming study by Scheffler and Nenchev (2024), which includes up to three free-form sense descriptions for each emoji collected from participants between 18 and 70 years of age.

We divided the participants at the median age into two age groups, yielding one younger group (28 or younger, $n = 74$) and one older group (29 or older, $n = 77$). The age distribution can be seen in Figure 16 (in Appendix B). Most participants were younger than 40 years of age (approximately 87%), resulting in a broader age range within the older group.

To measure the similarity between the age groups, one vector representation per emoji and age group was created by averaging the word embeddings of the collected emoji descriptions from Scheffler and Nenchev (2024). For comparability, the same *spaCy*-embeddings used by Scheffler and Nenchev (2024) were employed. We removed stop words using *NLTK* and calculated the cosine similarity between the resulting vectors for each emoji. The average cosine similarity between age groups was 0.83 ($SD = 0.09$). Figure 17 (in Appendix B) presents the similarity values for all emojis. Overall, most emojis have similar meanings across age groups, suggesting relatively stable interpretations independent of age. However, some emojis, such as *Skull* (💀), showed noticeably lower similarity values, indicating stronger differences in interpretation between age groups.

To examine whether age-based variation is associated with semantic change over time, we aggregated the overlap values we calculated into a single score for each emoji. For this purpose, we chose the difference between the maximum score (smallest overlap from year to year) and the mean overlap across all year pairs:

$$D(A) = max(A) - \bar{a}$$

where D denotes the difference between the maximum value in list A and the mean of all values in the list ($\bar{a}$). We were able to do this for all 73 emojis that were already available from 2015 onward. This measure was chosen to distinguish localized semantic shifts from generally unstable representations. Emojis with consistently low overlap values would also have comparatively low D-values, whereas emojis with temporary decreases in overlap would produce higher values. Figure 18 (in Appendix B) shows all values calculated in this manner sorted by emoji. Most emojis show small differences between their lowest overlap and the average of all overlaps, indicating a stable representation. The mean of all calculated values was 89.74 ($SD = 56.28$). This is consistent with the results by Robertson et al. (2021), who found that most emojis have fairly stable representations in English datasets. The emojis *Skull* (💀) ($D = 249.57$) and *Face with Medical Mask* (😷) ($D = 308.86$) showed the strongest indications of semantic change.

Finally, we tested the possible relationship between age-based (apparent time) meaning differences and (real time) semantic change using *Spearman's Rank Correlation*. We found a weak positive correlation between the age-based cosine distances (1 - cosine similarity) and D-values ($\rho = 0.2$, $p = 0.03$). The relationship is visualized in Fig. 3.

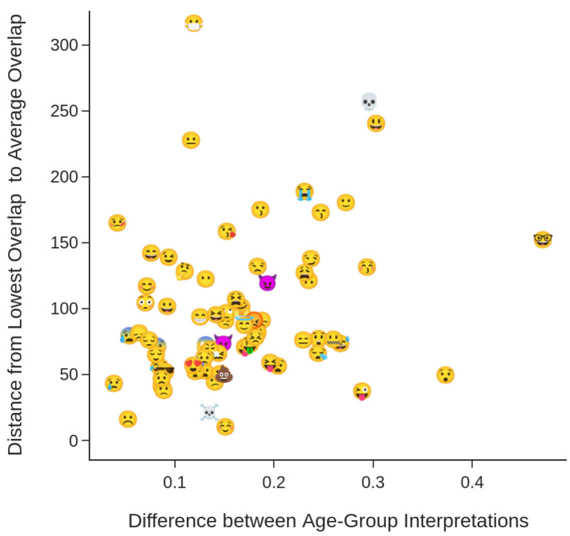


Figure 3: Visualization of the correlation of age-based description differences and indication for a meaning change for the 73 emojis available for the full examination period.

These findings suggest that age-based differences in emoji meanings may provide information about ongoing semantic change. However, age-related variation alone cannot explain all observed changes. Some emojis like *Face with Medical Mask* (😷) show a high indication for a meaning change, yet the emoji does not show vast differences in the interpretation between age groups. The following case studies therefore examine the four example emojis in greater detail to further investigate the relationship between semantic change and age-based variation.

10 Case Studies

10.1 Face with Tears of Joy

Face with Tears of Joy (😂) is the most widely used emoji (Scheffler and Nenchev, 2024). It also showed comparatively stable overlap values across consecutive years. To further examine its semantic stability, we calculated the overlap score directly between the embedding spaces trained on data from 2015 and 2022. Despite the seven-year gap, the emoji retained an overlap of 318 out of 1000 nearest neighbors, suggesting a stable semantic representation.

To qualitatively inspect the structure of the embeddings, we combined the 50 nearest neighbors from both embedding spaces and visualized them using *t-SNE* (Figure 4). Neighbors unique to 2015 are shown in yellow, neighbors unique to 2022 in blue, and overlapping neighbors in green. Doing so allows us to see how distinct and distant the lists of nearest neighbors are (Gonen et al., 2020).

Only eight neighbors appeared in both top 50 lists, resulting in relatively little direct overlap. Nevertheless, qualitative inspection suggests substantial semantic similarity between the neighborhoods. Many close neighbors in both years consist of orthographic variants of laughter expressions such as "haha", including lengthened spellings and variants such as "Muhaha" or "Bwahaha". While these forms differ lexically, they express similar meanings.

This finding highlights a limitation of neighbor overlap approaches when applied to social media data. Rapid changes in spelling conventions and expressive writing styles can reduce neighbor overlap without reflecting actual semantic change. Despite this, the merged neighbor lists still showed high average cosine similarities in both embedding spaces (0.63 for 2015 and 0.58 for 2022). The results suggest that *Face with Tears of Joy* remained semantically stable across the entire examined period despite substantial lexical variation in its contextual neighborhoods.

10.2 Smiling Face

Smiling Face (☺️) is among the most frequently used face emojis, similar to *Face with Tears of Joy*. However, unlike the latter, it does not show stable neighbor overlap across years. Instead, overlap between the nearest neighbors in the yearly embeddings remains consistently low.

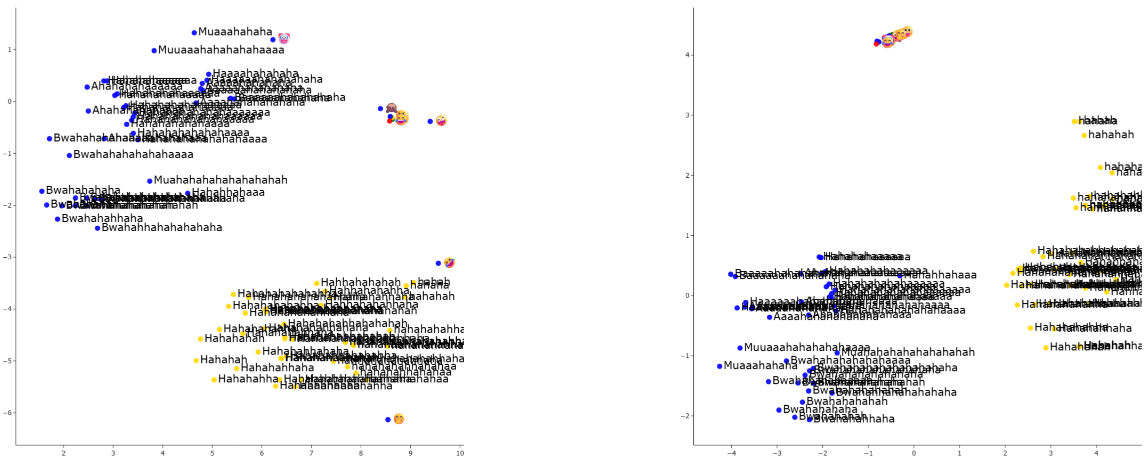
To further examine this instability, we extracted the 1000 nearest neighbors for 2015 and 2022 and visualized the 50 closest neighbors using *t-SNE* (Figure 5). No overlap was found in either the top 50 or top 1000 neighbor sets.

While the 2022 neighbors show a loose thematic association with university-related contexts, no consistent semantic pattern emerges for 2015. This suggests that *Smiling Face* is used in highly variable contexts without a stable dominant context. To explore this further, we manually inspected randomly selected tweets with this emoji (Table 4 in Appendix C). The emoji frequently co-occurs with heart emojis and appears in general social interaction contexts rather than topic-specific usage. This emoji has a strongly evaluative and expressive meaning (Grosz, 2026) and is thus poorly characterized by other context words. The consistently low neighbor overlap demonstrates that distributional representations of the kind used here have limited value for certain types of expressive phenomena, such as emojis like ☺️.

10.3 Face with Medical Mask

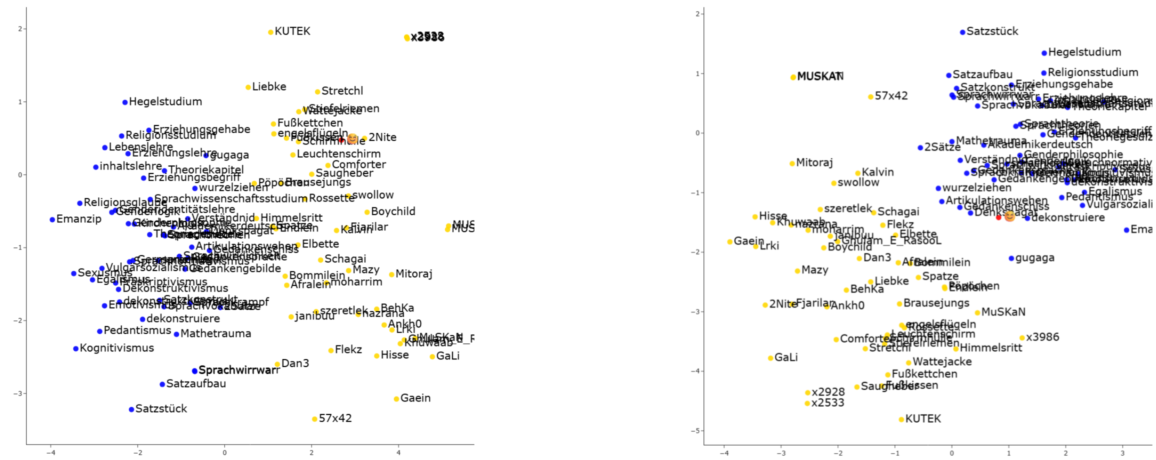
The emoji *Face with Medical Mask* (😷) shows a temporary decrease in overlap between the 1000 closest neighbors in 2019 and 2020. From 556 (2018–2019), the number drops to only 123 overlaps (2019–2020). This shift coincides with a strong increase in usage frequency, from 18,442 occurrences in 2019 to 135,730 in 2020 (+636%) in the corpus collected using the method described in Scheffler (2014). This change is likely related to the onset of the COVID-19 pandemic, which shifted the contextual usage of the emoji from general illness-related content toward pandemic-specific discourse.

A comparison of the 50 nearest neighbors in 2019 and 2020 (Figure 6) shows 11 overlapping items, all of which are emojis. While the 2019 neighbors are predominantly emojis, the 2020 neighborhood contains more lexical items referring to masks and public health measures.



(a) 2015 (b) 2022

Figure 4: T-SNE visualization of 50 nearest neighbors of *Face with Tears of Joy* in 2015 (yellow) and 2022 (blue). Overlapping neighbors are shown in green. Embeddings based on Twitter-Data from 2015 (left) and 2022 (right).



(a) 2015 (b) 2022

Figure 5: T-SNE visualization of 50 nearest neighbors of *Smiling Face* in 2015 (yellow) and 2022 (blue). Overlapping neighbors are shown in green. Embeddings based on Twitter-Data from 2015 (left) and 2022 (right).

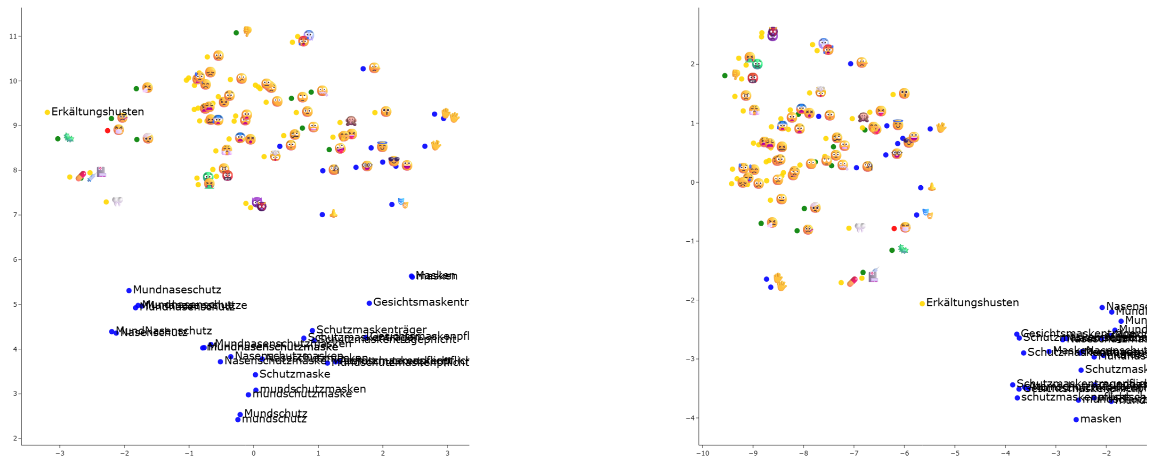
The nearest-neighbor structure, therefore, reflects the semantic shift from general illness to COVID-19-related meanings. This suggests a rapid, externally triggered meaning change affecting all user groups simultaneously, independent of age. It is thus not surprising that the emoji does not fit the *Apparent Time Hypothesis* (as operationalized in Section 9).

10.4 Skull

Similarly to *Face with Medical Mask*, *Skull* (💀) experiences a decrease in overlap between the lists of nearest neighbors. However, this decrease spans multiple years (2016–2019). This aligns with descriptions of a shift in usage among younger users, where the emoji increasingly represents laughter and irony instead of death (Austin, 2024).

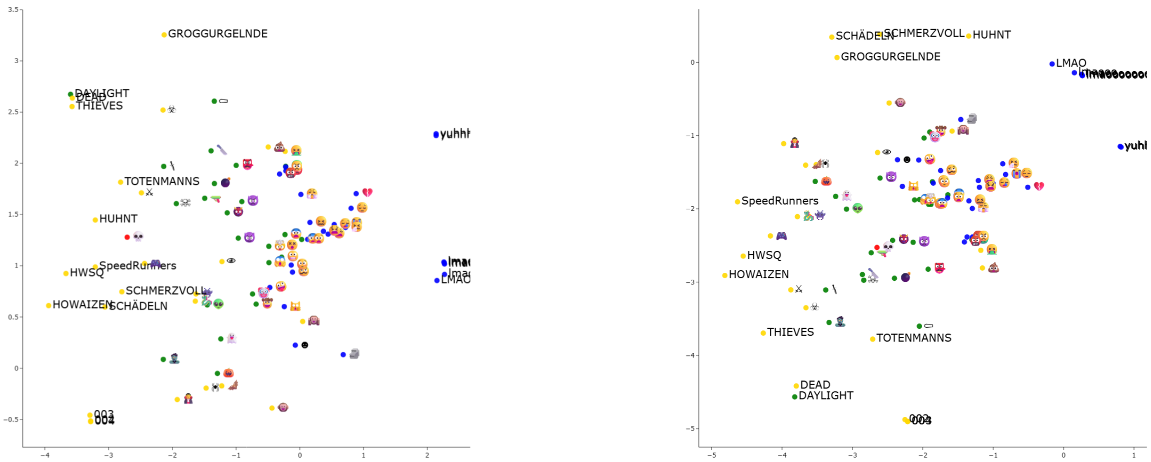
The lowest overlap occurs between 2018 and 2019 (131 among the top 1000 neighbors). These years were selected for closer inspection (Figure 7). Among the top 50 neighbors, 24 items overlap between the two years, suggesting a more stable meaning in the most central part of the neighborhood.

The nearest neighbors form a cloud from left (older meanings) to right (newer), without distinctly separated clusters. While in 2018 the semantic field of video games can be observed, for 2019 suddenly words such as "*LMAO*" appear. These mirror the increasing use of the emoji as a sign of laughter, described by Austin (2024). The emoji *Skull* (💀) can thus be seen as an example of a slow and socially driven meaning shift rather than a sudden contextual change. This slow meaning change matches the



(a) 2019 (b) 2020

Figure 6: T-SNE visualization of 50 nearest neighbors of *Face with Medical Mask* in 2019 (yellow) and 2020 (blue). Overlapping neighbors are shown in green. Embeddings based on Twitter-Data from 2019 (left) and 2020 (right).



(a) 2018 (b) 2019

Figure 7: T-SNE visualization of 50 nearest neighbors of *Skull* in 2018 (yellow) and 2019 (blue). Overlapping neighbors are shown in green. Embeddings based on Twitter-Data from 2018 (left) and 2019 (right).

Apparent Time Hypothesis as operationalized by the lower cosine similarities between the meanings that older vs. younger users show for this emoji (0.7) as compared to *Face with Medical Mask* (0.88) with the age-independent meaning change.

11 Conclusion

We presented the first investigation of semantic change in face emojis in German-language social media data. We tracked the meaning of 107 emojis in a dataset spanning from 2015 to 2022, using the method proposed by Gonen et al. (2020). We discovered several patterns of meaning overlap for emojis across the timespan. While most emojis were shown to have stable meanings, we identified both socially driven semantic change, which corresponds to *apparent time* meaning differences influenced by the age of speakers, as well as contextual meaning shifts occurring across all user groups. Furthermore, our findings show that even frequently used emojis that are interpreted similarly by users across age groups do not necessarily exhibit stable representations with computational semantic methods based on corpus distributions.

Overall, our study provides insights into the semantic change of face emojis in social media communication, highlights limitations of existing computational approaches for this type of data, and expands on research on the role of social and demographic factors on meaning shifts.

Limitations

New emojis are continuously added to the *Unicode* lists, limiting the number of testable emojis for certain years. Furthermore, newly introduced emojis tend to be used more infrequently. As a result, data availability for certain emojis is limited, especially with lesser spoken languages. A further limitation concerns the use of social media data, as platforms tend to be very age specific, reducing the (linguistic) diversity of the training data. Finally, the available data for examining differences between age groups is constrained. The norming study by Scheffler and Nenchev (2024) was not specifically conducted for comparison across age groups. Thus, most participants were between 18 and 40 years old, resulting in a fairly homogeneous age group with fewer linguistic differences.

Ethics Statement

All social media data used was collected from public sources and was solely processed in aggregated form. We are not aware of any other ethical issues affecting this work.

Acknowledgements

This study was partially funded by the Deutsche Forschungsgemeinschaft (DFG) through the Priority Program DFG SPP 2392 Visual Communication (ViCom), project EmDiCom (project number 501983851).

Calculations for this study were performed on the HPC cluster Elysium at Ruhr University Bochum, subsidized by the DFG (INST 213/1055-1).

References

Eyal Arviv and Oren Tsur. 2022. How to do things without words: Modeling semantic drift of emoji. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 4206–4211, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Steven Austin. 2024. What does the skull emoji really mean? a deep dive. *Marketing Scoop (online)*.

Guy Bailey. 2004. Real and apparent time. In J.K. Chambers, P. Trudgill, and N. Schilling-Estes, editors, *The Handbook of Language Variation and Change*, chapter 12, pages 312–332. John Wiley & Sons, Ltd.

Francesco Barbieri, Francesco Ronzano, and Horacio Saggion. 2016. What does this emoji mean? a vector space skip-gram model for Twitter emojis. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 3967–3972, Portorož, Slovenia. European Language Resources Association (ELRA).

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146.

Isabelle Boutet, Jean-Christophe Goulet-Pelletier, Eva Sutera, and Bozana Meinhardt-Injac. 2024. Are older adults adapting to new forms of communication? a study on emoji adoption across the adult lifespan. *Computers in Human Behavior Reports*, 13:100379.

Isabelle Boutet, Joëlle Guay, Justin Chamberland, Denis Cousineau, and Charles Collin. 2023. Emojis that work! incorporating visual cues from facial expressions in emojis can reduce ambiguous interpretations. *Computers in Human Behavior Reports*, 9:100251.

Hila Gonen, Ganesh Jawahar, Djamé Seddah, and Yoav Goldberg. 2020. Simple, interpretable and stable method for detecting words with usage change across corpora. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 538–555, Online. Association for Computational Linguistics.

Patrick G. Grosz. 2026. Expressivity and emojis. In Daniel Gutzmann and Katharina Turgay, editors, *The Oxford Handbook of Expressivity*. Oxford University Press.

Patrick Georg Grosz, Gabriel Greenberg, Christian De Leon, and Elsi Kaiser. 2023. A semantics of face emoji in discourse. *Linguistics and Philosophy*.

William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2016. Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1489–1501, Berlin, Germany. Association for Computational Linguistics.

Susan C. Herring and Ashley R. Dainas. 2020. Gender and age influences on interpretation of emoji functions. *ACM Transactions on Social Computing*, 3:1 – 26.

Maximilian Köper, Christian Scheible, and Sabine Schulte im Walde. 2015. Multilingual reliability and "semantic" structure of continuous word spaces. In *Proceedings of the 11th International Conference on Computational Semantics*, pages 40–45, London, UK. Association for Computational Linguistics.

Petra Kralj Novak, Jasmina Smailovic, Borut Sluban, and Igor Mozetic. 2015. Sentiment of emojis. *PloS one*, 10.

Alexander Robertson, Farhana Ferdousi Liza, Dong Nguyen, Barbara McGillivray, and Scott A. Hale. 2021. Semantic journeys: Quantifying change in emoji meaning from 2012-2018. *Workshop Proceedings of the 19th International AAAI Conference on Web and Social Media*.

Tatjana Scheffler. 2014. A German Twitter snapshot. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC)*.

Tatjana Scheffler and Ivan Nenchev. 2024. Affective, semantic, frequency, and descriptive norms for 107 face emojis. *Behavior Research Methods*, 56:8159–8180.

Shr Li Teh. 2025. Emoji semantic change and generational variation in digital communication. *figshare (Thesis)*.

A Embedding Evaluation Results

Category	**Example**	2015	2016	2017	2018	2019	2020	2021	2022
all		0.64	0.66	0.63	0.65	0.66	0.64	0.64	0.66
capital-common-countries	Athen Griechenland, Bagdad Irak	0.97	0.98	0.99	0.99	1	0.99	0.94	0.93
capital-world	Abuja Nigeria, Accra Ghana	0.69	0.70	0.65	0.67	0.70	0.63	0.65	0.69
currency	Algerien Dinar, Angola Kwanza	0.07	0.05	0.04	0.06	0.06	0.07	0.05	0.13
city-in-state	Chicago Illinois, Houston Texas	0.20	0.28	0.22	0.29	0.24	0.30	0.31	0.35
family	Junge Mädchen, Bruder Schwester	0.76	0.80	0.75	0.80	0.79	0.85	0.84	0.85
opposites	akzeptabel inakzeptabel, wissend unwissend	0.42	0.46	0.47	0.45	0.41	0.45	0.47	0.47
comparative	schlecht schlechter, beliebt beliebter	0.89	0.89	0.90	0.90	0.90	0.90	0.90	0.90
superlative	schlecht schlechteste, schwer schwerste	0.62	0.62	0.64	0.63	0.60	0.62	0.57	0.59
present-participle	codieren codierend, tanzen tanzend	0.60	0.60	0.60	0.56	0.59	0.57	0.61	0.60
nationality-adjective	Albanien albanisch, Argentinien argentinisch	0.82	0.85	0.83	0.86	0.86	0.82	0.82	0.84
past-tense	tanzen tanzte, sinken sinkte	0.75	0.73	0.71	0.73	0.79	0.74	0.71	0.74
plural	Banane Bananen, Vogel Vögel	0.81	0.85	0.80	0.85	0.85	0.87	0.84	0.76
plural-verbs	verringern verringert, beschreiben beschreibt	0.99	0.98	0.98	0.99	0.99	0.99	0.98	0.98

Table 3: Results of the Analogy Task (accuracy) by Köper et al. (2015) sorted by year. Row *all* shows results if tests are run together.

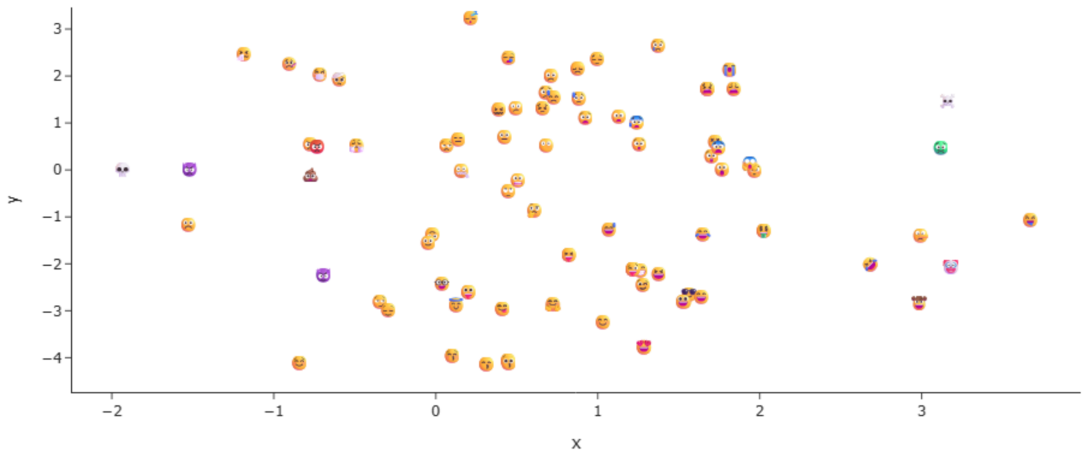


Figure 8: T-SNE visualization of face emoji embeddings (n=80) reduced to two dimensions based on the model for the year 2016.

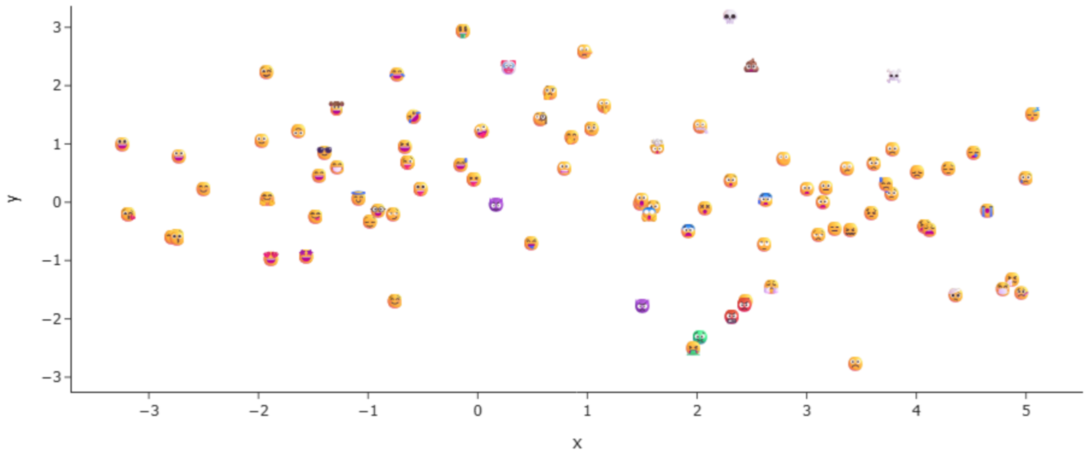


Figure 9: T-SNE visualization of face emoji embeddings (n=89) reduced to two dimensions based on the model for the year 2017.

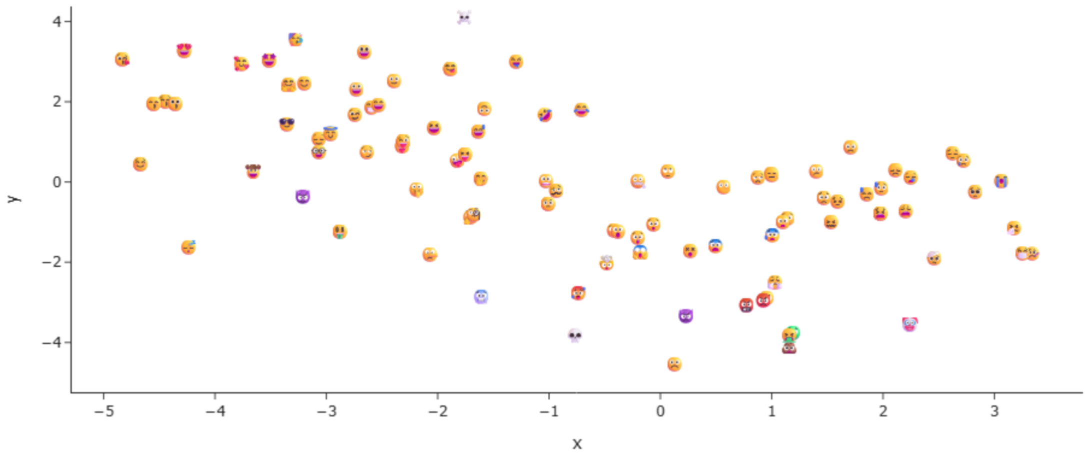


Figure 10: T-SNE visualization of face emoji embeddings (n=95) reduced to two dimensions based on the model for the year 2018.

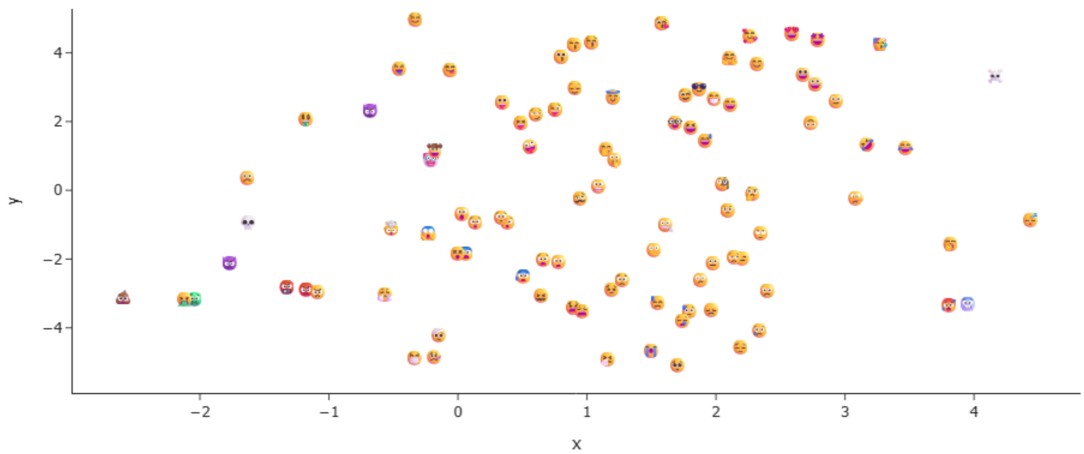


Figure 11: T-SNE visualization of face emoji embeddings (n=96) reduced to two dimensions based on the model for the year 2019.

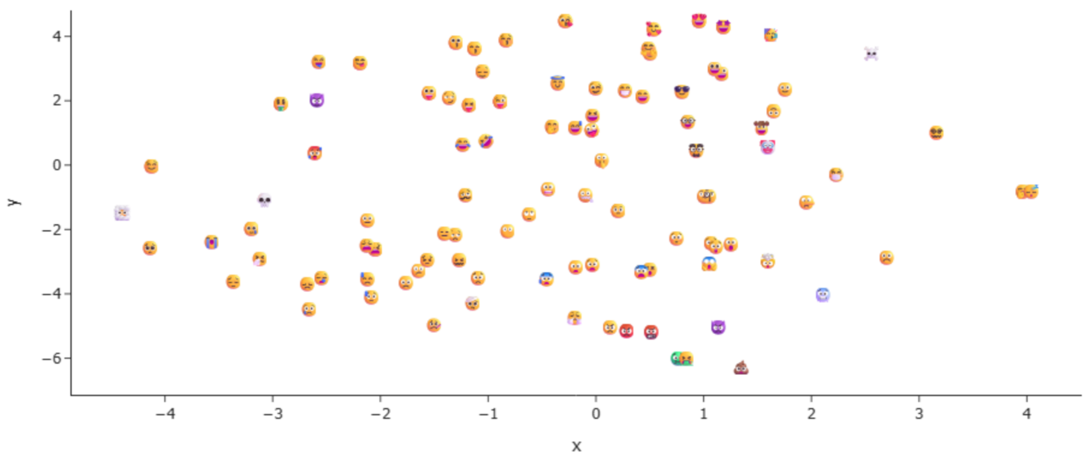


Figure 12: T-SNE visualization of face emoji embeddings (n=100) reduced to two dimensions based on the model for the year 2020.

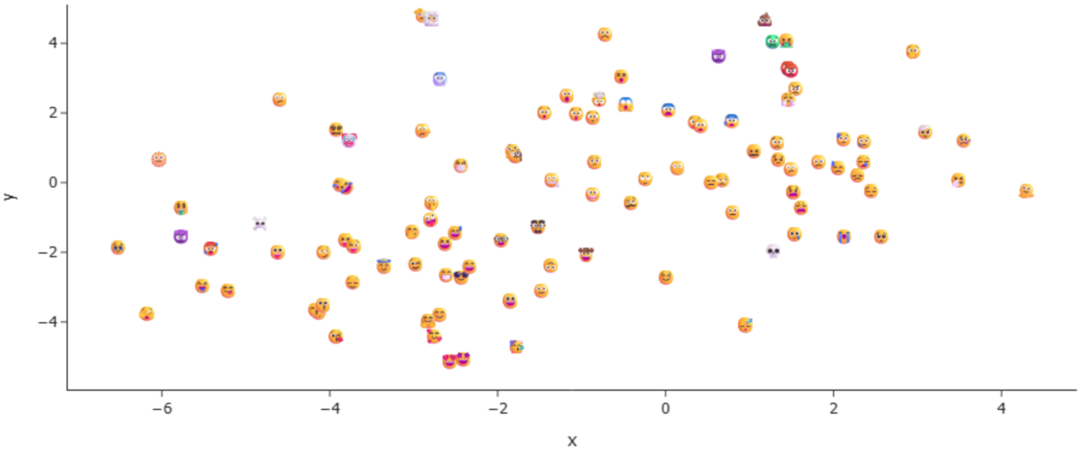


Figure 13: T-SNE visualization of face emoji embeddings (n=107) reduced to two dimensions based on the model for the year 2021.

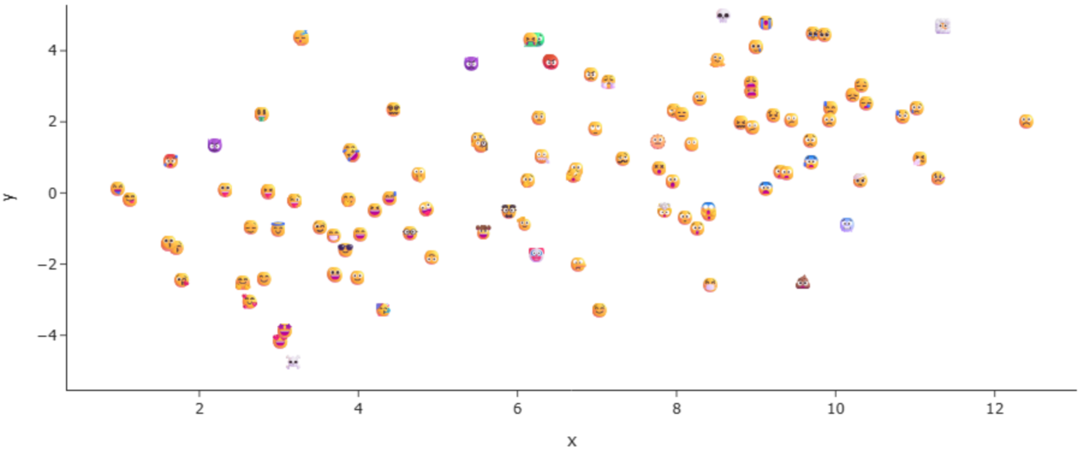


Figure 14: T-SNE visualization of face emoji embeddings (n=107) reduced to two dimensions based on the model for the year 2022.

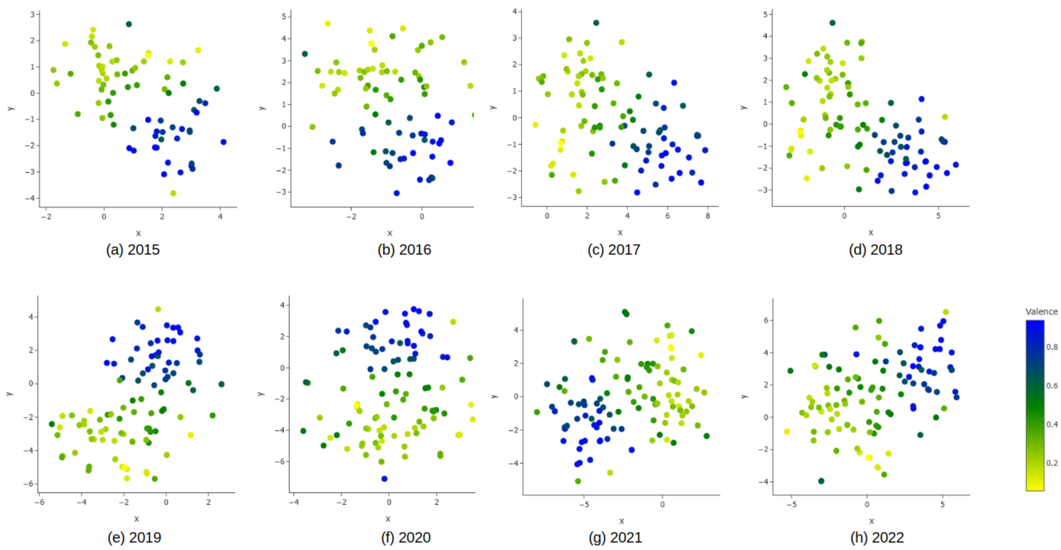


Figure 15: T-SNE visualization of face emojis for the years 2015 (a) to 2022 (h) with valence values based on Scheffler and Nenchev (2024). A positive valence has a high score (close to 1) and is shown by a blue marker, a negative valence score (close to 0) by a yellow marker. Emojis were only plotted if they were already available in the *Unicode*-lists during that year. Number of plotted emojis ranges from n=73 (2015) to n=107 (2021 and after).

B Age-Based Examinations

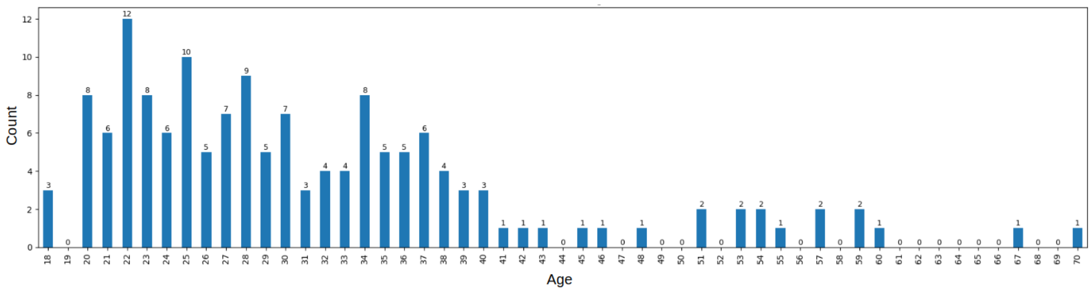


Figure 16: Age distribution of participants in the norming study conducted by Scheffler and Nenchev (2024).

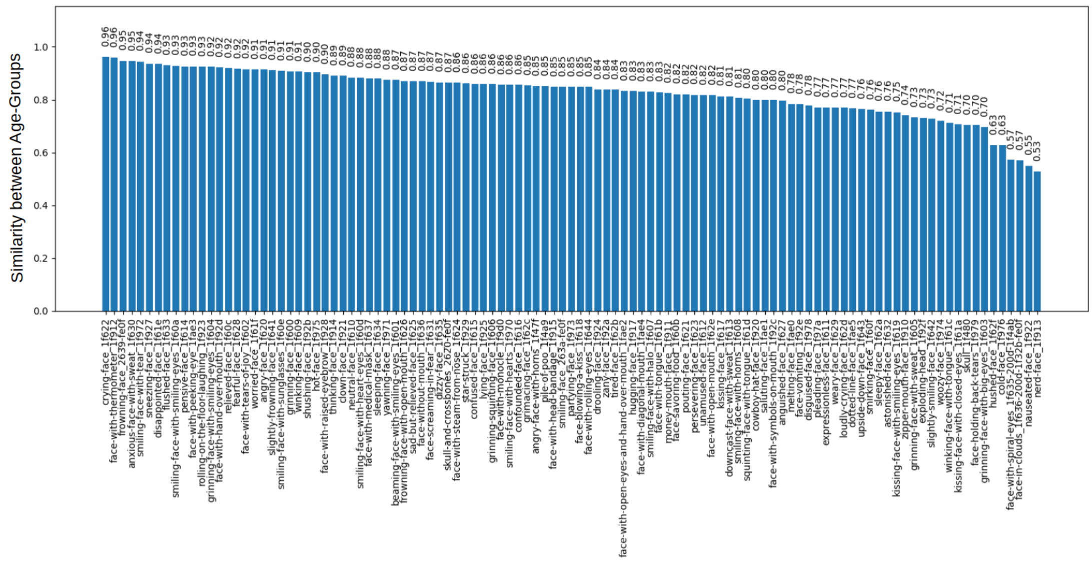


Figure 17: Similarity of emoji descriptions collected by Scheffler and Nenchev (2024) between the two examined age groups described in Section 9.

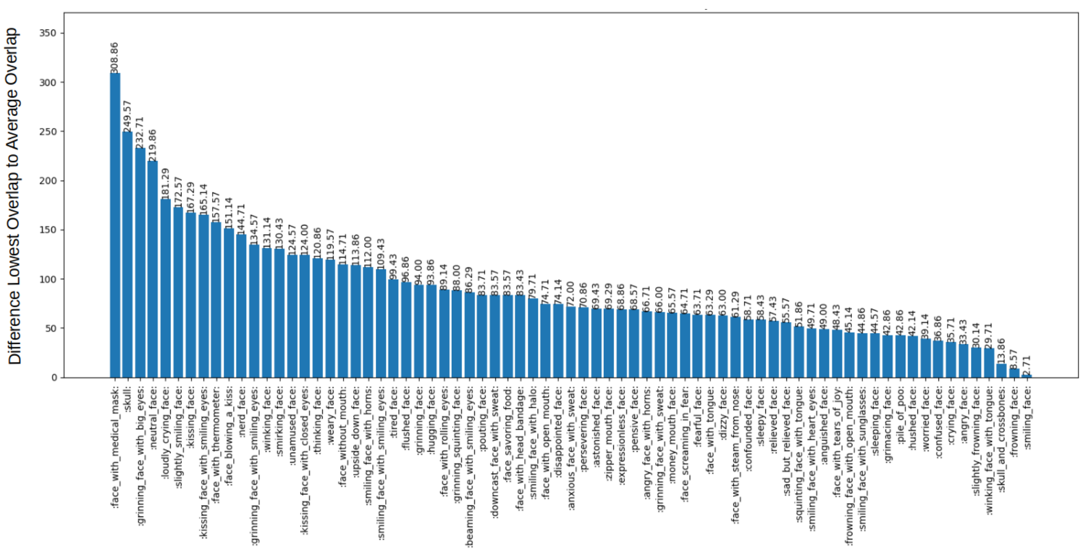


Figure 18: Differences between the lowest overlap and the average of all overlaps for each examined emoji (as described in Section 9).

C Corpus Examples

Year	Example Tweets
2015	1. @[MENTION] Jo, fürs video schauen reicht es vollkommen aus ☺️ 2. @[MENTION] Fröhlichen Mittwochmorgen auch dir.☺️ 3. Aufgewacht und meine ganze Startseite ist voll mit Tweets von dem @[MENTION] so wacht man doch gerne auf! 😂 👍 Guuuuuten Morgen 💖☺️ 4. zum Glück wartet im Büro scho 'Tante #Jura uf mi☺️ #kaffi 5. @[MENTION] ich liebe euch so sehr habe extra diesen Account nur für euch ! 😍❤️ um euch zu unterstützen ❤️ 💋 hoffe ihr schätzt es ☺️❤️
2022	1. @[MENTION] Jeder braucht Ziele auf die er hinarbeiten kann ☺️ 2. @[MENTION] @[MENTION] Bei uns auch.☺️ 3. @[MENTION] Dann hoffe ich, dass etwas Erinnerung geweckt werden konnte. 😉☺️ 4. @[MENTION] @[MENTION] Hallo, Dankeschön ☺️ Er ist ein Dogo Argentinio-Dogge Mischling und ist sieben Monate alt. Mein ganzer Stolz! ❤️ 5. @[MENTION] Von mir immer hübsche ☺️

Table 4: Example Tweets containing *Smiling Face* (☺️) randomly selected from the corpus collected using the method described by Scheffler (2014).

D Overview Overlap Scores

Emoji	Overlap Scores	Available	max	min	mean
:angry_face:	-471, -422, -426, -434, -405, -471, -440	7	-405	-471	-438.428571
:angry_face_with_horns:	-459, -337, -386, -396, -369, -432, -447	7	-337	-459	-403.714286
:anguished_face:	-551, -450, -515, -454, -477, -491, -555	7	-450	-555	-499.000000
:anxious_face_with_sweat:	-562, -471, -494, -545, -545, -606, -578	7	-471	-606	-543.000000
:astonished_face:	-493, -410, -482, -495, -511, -537, -407	7	-407	-537	-476.428571
:beaming_face_with_smiling_eyes:	-557, -444, -555, -516, -471, -574, -595	7	-444	-595	-530.285714
:clown_face:	-416, -387, -139, -524, -549, -600	6	-139	-600	-435.833333
:cold_face:	-544, -626, -603, -625	4	-544	-626	-599.500000
:confounded_face:	-559, -488, -530, -549, -515, -591, -595	7	-488	-595	-546.714286
:confused_face:	-550, -501, -541, -533, -498, -547, -574	7	-498	-574	-534.857143
:cowboy_hat_face:	-527, -400, -344, -443, -598, -574	6	-344	-598	-481.000000
:crying_face:	-481, -447, -494, -457, -459, -528, -513	7	-447	-528	-482.714286
:disappointed_face:	-524, -484, -540, -517, -409, -447, -461	7	-409	-540	-483.142857
:disguised_face:	-430, -507	2	-430	-507	-468.500000
:dizzy_face:	-499, -441, -475, -536, -524, -457, -596	7	-441	-596	-504.000000
:dotted_line_face:	0	1	0	0	0.000000
:downcast_face_with_sweat:	-538, -472, -550, -527, -501, -413, -475	7	-413	-550	-496.571429
:drooling_face:	-264, -435, -376, -440, -494, -494	6	-264	-494	-417.166667
:exploding_head:	-424, -539, -511, -559, -601	5	-424	-601	-526.800000
:expressionless_face:	-531, -468, -544, -539, -488, -590, -598	7	-468	-598	-536.857143
:face_blowing_a_kiss:	-492, -390, -295, -318, -516, -554, -558	7	-295	-558	-446.142857
:face_holding_back_tears:	0	1	0	0	0.000000
:face_in_clouds:	-68, -50	2	-50	-68	-59.000000
:face_savoring_food:	-494, -396, -465, -509, -483, -509, -501	7	-396	-509	-479.571429
:face_screaming_in_fear:	-527, -482, -477, -531, -571, -589, -615	7	-477	-615	-541.714286
:face_vomiting:	-488, -541, -577, -549, -423	5	-423	-577	-515.600000
:face_with_diagonal_mouth:	0	1	0	0	0.000000
:face_with_hand_over_mouth:	-444, -413, -503, -575, -442	5	-413	-575	-475.400000
:face_with_head_bandage:	-532, -458, -584, -515, -429, -553, -516	7	-429	-584	-512.428571
:face_with_medical_mask:	-538, -465, -536, -556, -123, -423, -382	7	-123	-556	-431.857143
:face_with_monocle:	-583, -521, -525, -545, -527	5	-521	-583	-540.200000
:f_w_op_eyes_n_hand_ov_mouth:	0	1	0	0	0.000000
:face_with_open_mouth:	-473, -462, -515, -500, -526, -409, -501	7	-409	-526	-483.714286
:face_with_peeking_eye:	0	1	0	0	0.000000
:face_with_raised_eyebrow:	-454, -427, -429, -447, -458	5	-427	-458	-443.000000
:face_with_rolling_eyes:	-402, -380, -417, -508, -483, -570, -524	7	-380	-570	-469.142857
:face_with_spiral_eyes:	-8, -7	2	-7	-8	-7.500000
:face_with_steam_from_nose:	-524, -445, -443, -491, -483, -583, -561	7	-443	-583	-504.285714
:face_with_symbols_on_mouth:	-379, -426, -435, -505, -399	5	-379	-505	-428.800000
:face_with_tears_of_joy:	-669, -581, -622, -647, -576, -616, -660	7	-576	-669	-624.428571
:face_with_thermometer:	-564, -558, -581, -609, -351, -442, -455	7	-351	-609	-508.571429
:face_with_tongue:	-489, -383, -456, -464, -399, -468, -465	7	-383	-489	-446.285714
:face_without_mouth:	-458, -371, -489, -478, -486, -528, -590	7	-371	-590	-485.714286
:fearful_face:	-494, -470, -490, -525, -566, -587, -604	7	-470	-604	-533.714286
:flushed_face:	-461, -401, -496, -543, -537, -516, -531	7	-401	-543	-497.857143
:frowning_face:	-12, -8, -10, -4, -7, -1, -25	7	-1	-25	-9.571429
:frowning_face_with_open_mouth:	-547, -459, -490, -479, -440, -455, -526	7	-440	-547	-485.142857
:grimacing_face:	-526, -445, -429, -444, -462, -513, -484	7	-429	-526	-471.857143
:grinning_face:	-578, -379, -473, -537, -412, -483, -449	7	-379	-578	-473.000000
:grinning_face_with_big_eyes:	-606, -190, -279, -462, -471, -451, -500	7	-190	-606	-422.714286
:grinning_face_with_smiling_eyes:	-379, -465, -598, -549, -492, -562, -550	7	-379	-598	-513.571429
:grinning_face_with_sweat:	-616, -535, -544, -500, -510, -460, -517	7	-460	-616	-526.000000
:grinning_squinting_face:	-575, -486, -536, -562, -448, -571, -574	7	-448	-575	-536.000000
:hot_face:	-404, -382, -434, -444	4	-382	-444	-416.000000
:hugging_face:	-551, -454, -565, -563, -530, -588, -584	7	-454	-588	-547.857143
:hushed_face:	-561, -451, -477, -458, -482, -513, -510	7	-451	-561	-493.142857
:kissing_face:	-458, -452, -555, -398, -258, -413, -443	7	-258	-555	-425.285714
:kissing_face_with_closed_eyes:	-513, -375, -549, -538, -443, -544, -531	7	-375	-549	-499.000000
:kissing_face_with_smiling_eyes:	-524, -447, -501, -495, -291, -481, -454	7	-291	-524	-456.142857
:loudly_crying_face:	-348, -527, -558, -542, -280, -450, -524	7	-280	-558	-461.285714
:lying_face:	-172, -404, -442, -411, -406, -338	6	-172	-442	-362.166667
:melting_face:	0	1	0	0	0.000000
:money_mouth_face:	-415, -367, -389, -413, -405, -522, -517	7	-367	-522	-432.571429
:nauseated_face:	-457, -553, -561, -591, -625, -589	6	-457	-625	-562.666667
:nerd_face:	-484, -431, -476, -602, -667, -675, -695	7	-431	-695	-575.714286

Emoji	Overlap Scores	Available	max	min	mean
:neutral_face:	-529, -379, -523, -481, -242, -559, -520	7	-242	-559	-461.857143
:partying_face:	-460, -547, -584, -595	4	-460	-595	-546.500000
:pensive_face:	-418, -447, -367, -427, -388, -500, -502	7	-367	-502	-435.571429
:persevering_face:	-582, -496, -551, -558, -546, -622, -613	7	-496	-622	-566.857143
:pile_of_poo:	-450, -417, -476, -474, -484, -467, -451	7	-417	-484	-459.857143
:pleading_face:	-301, -444, -468, -550	4	-301	-550	-440.750000
:pouting_face:	-477, -386, -437, -425, -398, -437, -329	7	-329	-477	-412.714286
:relieved_face:	-466, -422, -502, -469, -447, -517, -533	7	-422	-533	-479.428571
:rolling_on_the_floor_laughing:	-433, -641, -669, -612, -654, -672	6	-433	-672	-613.500000
:sad_but_relieved_face:	-543, -472, -539, -530, -544, -550, -515	7	-472	-550	-527.571429
:saluting_face:	0	1	0	0	0.000000
:shushing_face:	-484, -527, -558, -617, -623	5	-484	-623	-561.800000
:skull:	-477, -193, -265, -131, -350, -600, -648	7	-131	-648	-380.571429
:skull_and_crossbones:	-8, -3, -65, -6, -3, -12, 0	7	0	-65	-13.857143
:sleeping_face:	-608, -568, -627, -606, -603, -651, -625	7	-568	-651	-612.571429
:sleepy_face:	-528, -444, -480, -521, -471, -536, -537	7	-444	-537	-502.428571
:slightly_frowning_face:	-509, -471, -516, -467, -470, -519, -528	7	-467	-528	-497.142857
:slightly_smiling_face:	-509, -481, -560, -497, -297, -524, -419	7	-297	-560	-469.571429
:smiling_face:	-7, -4, -1, -1, 0, -5, -1	7	0	-7	-2.714286
:smiling_face_with_halo:	-447, -468, -517, -548, -535, -581, -591	7	-447	-591	-526.714286
:smiling_face_with_heart_eyes:	-358, -417, -385, -409, -485, -413, -387	7	-358	-485	-407.714286
:smiling_face_with_hearts:	-498, -524, -562, -525	4	-498	-562	-527.250000
:smiling_face_with_horns:	-428, -406, -522, -350, -433, -544, -551	7	-350	-551	-462.000000
:smiling_face_with_smiling_eyes:	-485, -403, -507, -540, -514, -554, -584	7	-403	-584	-512.428571
:smiling_face_with_sunglasses:	-570, -537, -591, -583, -593, -633, -566	7	-537	-633	-581.857143
:smiling_face_with_tear:	-569, -587	2	-569	-587	-578.000000
:smirking_face:	-447, -359, -442, -467, -529, -586, -596	7	-359	-596	-489.428571
:sneezing_face:	-410, -433, -406, -203, -430, -214	6	-203	-433	-349.333333
:squinting_face_with_tongue:	-484, -410, -434, -450, -454, -504, -497	7	-410	-504	-461.857143
:star_struck:	-520, -493, -538, -554, -449	5	-449	-554	-510.800000
:thinking_face:	-397, -261, -365, -431, -394, -382, -443	7	-261	-443	-381.857143
:tired_face:	-461, -459, -551, -597, -590, -631, -620	7	-459	-631	-558.428571
:unamused_face:	-543, -393, -539, -523, -505, -572, -548	7	-393	-572	-517.571429
:upside_down_face:	-493, -481, -485, -499, -351, -552, -393	7	-351	-552	-464.857143
:weary_face:	-460, -429, -545, -599, -587, -595, -625	7	-429	-625	-548.571429
:winking_face:	-603, -372, -413, -551, -494, -532, -557	7	-372	-603	-503.142857
:winking_face_with_tongue:	-566, -516, -548, -528, -516, -510, -496	7	-496	-566	-525.714286
:woozy_face:	-305, -402, -616, -494	4	-305	-616	-454.250000
:worried_face:	-541, -467, -511, -511, -461, -484, -526	7	-461	-541	-500.142857
:yawning_face:	-265, -526, -521	3	-265	-526	-437.333333
:zany_face:	-349, -407, -421, -507, -470	5	-349	-507	-430.800000
:zipper_mouth_face:	-450, -392, -451, -396, -365, -470, -516	7	-365	-516	-434.285714

Table 6: Overview of all overlap scores based on self-trained embeddings using the method described by Gonen et al. (2020) sorted by emoji. Higher values (closer to 0) show little overlap between the closest neighbors of two year-based vector spaces, indicating a potential meaning shift or unstable representation.