



SADNESS-RELATED EXPRESSIONS IN DANISH AND GERMAN: A CORPUS-ASSISTED NSM-ANALYSIS

Katalin Fenyvesi
University College Absalon,
University of Southern Denmark
kafe@pha.dk

Eckhard Bick
University of Southern Denmark

Klaus Geyer
University of Southern Denmark

Abstract

The study explores sadness-related expressions in two typologically closely related languages in the natural semantic metalanguage (NSM) framework. A systematic corpus enquiry revealed the syntactic patterns and helped to identify the most frequent head-nouns of a number of Danish and German sadness-related expressions. German *traurig*, for instance, has a distribution similar to that of Danish *sørgelig* with semiotic products and clauses as subjects. However, when used with human subjects, its distribution aligns with the Danish multi-word expression *ked af det*. Semantic consultations conducted about the use of the most salient sadness adjectives with some speakers of Danish and German revealed fine-grained differences between German *traurig* and *trist* and Danish *ked af det* and *trist* respectively. Thus, when used with a human headword, Danish *trist* is more trait-like while *ked af det* is more state-like. The concept of sadness-related emotions in Danish and German is discussed, followed by a methodological discussion about the combinability of a quantitative corpus approach, a qualitative semantic consultation approach and NSM explications. Corpus inquiry was used to chart the adjectives' polysemy, and as a method for creating the NSM explications, consultation data were used.

Keywords: sadness-related expressions, corpus linguistics, lexical semantics, Danish, German, natural semantic metalanguage (NSM)

1. Introduction

Certain concepts are often believed to be universal, but a closer semantic analysis suggests that they are not. The emotion sadness is, according to Kemper (1987), one of four basic psychological

emotions: fear, anger, sadness, and satisfaction. One of his reasons for calling these emotions basic was the assumption that they were universally found in all cultures. Later research on emotions (Ekman et al. 2013) described six basic emotions, based on facial expressions: happiness, surprise, fear, sadness, anger, and disgust combined with contempt, claiming these to be universal. In psychological research on emotions, the assumption that emotions are universal is often left out of consideration. Nevertheless, emotions are usually described in English terms, although different languages use very different terms for denoting emotions and the categorization of emotions or emotional concepts differs to a huge degree between cultures (Russell 1991). In psychological research, English concepts are often unconsciously taken for granted as universals in the description of emotions. In contrast, the emotion “sadness” is described as a culture-specific concept by others, because an expression for sadness is apparently lacking in some languages. In many Pacific languages, for example Tahitian, there is no expression for “sad” (Levy 1984; Goddard 2007). Sadness-related expressions have been compared through relatively distant languages, such as English and Chinese (Zhang 2016) and among Indo-European languages, such as English and Russian (Wierzbicka 1999). Sadness-related expressions have been described and compared through languages within different linguistic theories and methodologies such as corpus linguistics (Zhang 2016) and the natural semantic metalanguage (NSM) approach (Wierzbicka 1999). Wierzbicka (2009) argues that emotion terms in many languages cover very different concepts than the English terms, thus English is not a neutral scientific language for the description of emotions. Therefore, in commitment to non-Anglocentric analysis, the NSM approach should be used for the description of often ethnocentric emotional concepts. This paper follows Wierzbicka’s suggestion and compares sadness-related expressions in Danish and German, two genetically and typologically closely related neighbouring Germanic languages in Europe, which have not been studied from a comparative point of view so far. A further invention of this study is the combination of systematic quantitative corpus analysis and semantic consultations with the NSM approach.

2. Theory and previous research

2.1. Sadness-related expressions in German and Danish

Danish and German have different ways of expressing sadness. The German adjective *traurig* ‘sad’¹ can in many contexts be equivalent with the Danish adjectival expression *ked af det* ‘sad’ if it refers to a person. In contexts where *traurig* refers to the object causing the emotion of sadness, on the other hand, it can be equivalent with the Danish adjective *trist* ‘sad’. In some very specific contexts, the Danish adjective *kedelig* ‘boring, embarrassing or flat’ can also be equivalent with *traurig*. The German onomasiological dictionary by Dornseiff (2012), and its Danish equivalent by Nimb (2014), contains a number of frequently and rarely used adjectives, verbs, and nouns in the semantic field of sadness-related emotions. Since we in the corpus approach also want to examine syntactic patterns of the words but had to restrict the investigation to a manageable number of words, we chose to focus on frequently used adjectives that seemed to be central within this semantic field.

It has been suggested that expressions consisting of a single lexical item are easier to process, rather than a phrase or a clause (Slobin 2006). It is important to note that Danish is lacking a single adjective similar to English *sad* or German *traurig*. The concept is expressed by the construction *ked af det* which is a phrase rather than a word, since it can occur in a changed order with a dummy subject ‘*det*’: *det er jeg ked af* ‘this makes me sad’ literally ‘that am I sad of’. On the other hand, *ked af det* is syntactically restricted since it cannot be used as a prenominal attribute, unlike, for example, English *well-off* or *well-to-do*. Possibly this is due to a closeness condition on prenominal adjectival inflection. Thus, *forældrene er kede af det* ‘the parents are sad’ does allow plural inflection (*kede*) in predicative position, but the interfering particles *af* and *det* seem to block it in prenominal position: **de kede af det forældre* ‘the sad parents’.

2.2. NSM and corpus studies in lexical semantics

One of the central purpose of NSM explications is to differentiate between the meaning of near synonymous lexical items which are often defined in terms of each other in dictionaries. The decomposition into universal primes and reductive paraphrases is perfectly suited for the description of meaning differences between

¹ English translations are only translation equivalents and cannot be regarded as full semantic equivalents.

synonyms, since explications avoid circularity. NSM explications are also often used for cross-linguistic comparisons of near synonyms in different languages where the NSM paraphrases serve as a *tertium comparationis* (Ye 2013). NSM linguists often focus on the meaning analysis of language- and culture-specific words (Wierzbicka 1997; Levisen & Waters 2017. For a detailed list of state-of-the-art publications see <https://nsm-approach.net>). For example, the meaning of English *sadness* and the two closest counterparts in Russian, *pečal'* and *grust'*, are compared by Wierzbicka (1999). Goddard and Karlsson (2008) compared the meaning of the English verb *think* with the related Swedish verbs *tycka*, *tänka*, and *tro*. Levisen (2014) described the meaning difference between the English noun *happiness* and the Danish noun *lykke*, emphasizing that while the English concept of happiness includes elements like always smiling, Danish *lykke* refers either to an everyday wellbeing if referring to a long lasting, gradable state, or to a momentary, non-gradable state of extremely positive emotions. In his later work, Levisen (2018) explored the Danish adjectives *pæn* 'good-looking', *flot* 'attractive', *dejlig* 'nice' and *lækker* 'sexy'. Inspired by Levisen (2012), an NSM explication of the German noun *Gemütlichkeit* 'coziness' has been proposed by Geyer (2015) but no comparisons between Danish and German have been made in an NSM approach.

NSM paraphrases are also suited for identifying and distinguishing between the polysemous meanings of words. In this case, the reductive paraphrases of polysemous word meanings are characterized by overlap (Goddard 2000). Goddard (2000) analyzes the meaning of seemingly polysemous words, such as the verb *love*. He points out that the verb can occur in different syntactic frames, such as "person-X loves person-Y", or "person-X loves object-Y" which would be analyzed by two different NSM explications. The expression *in love (with)* would also be analyzed as a different explication because it has a much more specific meaning. On the other hand, according to Goddard, two uses of the same word are analyzed as one common meaning, if it is possible to find a common paraphrase for them, which is narrow enough to distinguish it from other related words.

Corpus-based research in lexical semantics is a usage-based approach enquiring about language use in context. It assumes that the contexts a word occurs in are indicative of its meaning. The accessibility of large amount of naturally occurring language usage

makes it possible to statistically analyze big data. The approach introduced by Sinclair (1991) describes the meaning of words as a function of their typical lexical environment (collocations) and syntactic patterns (colligations). A more recent, purely statistical approach are word embeddings, distributed representations (Mikolov et al. 2013; Heylen et al. 2015) where the semantics of a word is represented by a vector in a multi-dimensional space, where co-occurring content words, either adjacent or anywhere in the sentence, are used as dimensions. Cosine distances between vectors can then be used to estimate word similarity and to lump words into clusters automatically, rather than assigning them linguistically motivated categories by hand. Another corpus approach in lexical semantics is dependency-based, where the syntactic structure of the sentences must be marked in the corpora with a parser (Bick 2009).

NSM linguists often use introspection and text corpora as data sources (Wierzbicka 1998, 1999; Levisen 2018) without doing a systematic corpus analysis. Corpus linguists sometimes refer to NSM meaning decomposition (i.e. Zhang (2016)) but do not use the NSM approach in combination with corpus studies. NSM linguists, such as Tissari and colleagues (2019) use a corpus-assisted approach as a basis for formulating NSM paraphrases.

3. Research questions

The first aim of this paper is to analyze the similarities and differences between the semantic structure of some Danish and German adjectives and adjectival expressions expressing sadness-related emotions. The second aim is to find the possibilities of combining systematic quantitative corpus analysis, qualitative semantic consultations, and NSM paraphrases in the description of sadness-related expressions in two languages. The study seeks to answer the following research questions:

- RQ1 How do expressions for sadness-related emotions differ conceptually, lexically, and/or grammatically in Danish and German, and what do they have in common?
- RQ2 How can systematic corpus-analysis, qualitative semantic consultations, and NSM explications be combined in comparing sadness-related emotion words in these two languages?

4. Methods

As a starting point for finding the relevant sadness-adjectives in the two languages we used mono- and bilingual dictionaries. For both languages, the bilingual dictionaries *Gyldendals Røde Ordbog* and *Gyldendals Store Ordbog* were used. First, we looked up the Danish correspondences of the German adjective *traurig*. In the second step we looked at the back translations of these adjectives from Danish to German. The third step was to look up the above Danish adjectives in *Den Danske Ordbog* (DDO) and the above German adjectives in the Duden.de dictionary. The list was extended by adjectives that these dictionaries mentioned as synonyms or keywords in the definitions of the above. We chose those adjectives that were used most frequently in the dictionary definitions because we assumed that these were the most basic adjectives.

In order to complement the hand-crafted, monolingual dictionary definitions with an empirical perspective, two different research methods, one corpus-based and one semantic consultation-based, were used to investigate the semantics of the involved adjectives. In the first, quantitative, approach, we supplemented ordinary corpus inspection and word counts with data from DeepDict, a special dependency-based correlation analysis (Bick 2009), to shed light on syntactic constraints—that is, predicative and attributive use of the adjectives, as well as semantic co-occurrence patterns and contextual usage of the adjectives. In lexicographical terms, DeepDict is a relational dictionary that tries to constrain the semantic space occupied by a given word (lemma1 + POS) through the strength of syntactically linked co-occurrences. The method counts the frequency of head-dependent pairs ‘ $a \rightarrow b$ ’ in a mixed corpus and computes mutual information (co-occurrence strength) by normalizing pair frequency against the product of the stand-alone frequencies of words a and b in the corpus as a whole. This way, a ranked list of typical syntactic collocates can be created—for example, for verbs, typical subjects and objects; for adjectives, typical head nouns. For instance, DeepDict can implicitly “explain” the semantic difference between *big*, *high*, *large*, and *tall* (all translatable into German as *groß*) by providing typical head nouns, and it can also identify fixed expressions such as *big bang*. Data from *Korpus90/2000/2010* (ca. 86 M words) and from the *Leipzig Internet Corpus* (ca. 47 M words) were used.

Secondly, we performed explorative semantic consultations with native speaker informants of Danish and German. Our informants were a group of eight adult native speakers of Danish and four native speakers of German. The semantic consultations were conducted in the mother tongue of the informants. Informants were asked which adjectives they used in situations where something bad happened to somebody. Which nouns can these adjectives be combined with? What is the difference between the meaning of these adjectives? In which situations can they be used or not be used? The semantic consultations gave insight into native speakers' language use and intuitions about the meaning of the adjectives.

Finally, we wrote NSM paraphrases (Wierzbicka 1999) to describe and compare the meaning of the two most central Danish adjectives, *ked af det* and *trist*, and the two most central German adjectives, *traurig* and *trist*. We focused on these few adjectives because of space restrictions. These adjectives were chosen because they were used with the broadest range of headwords according to the corpus data. Corpus data also revealed that these adjectives had different meanings when used with human and non-human headwords. Thus, the analysis of the corpus data was helpful for deciding which meanings needed separate explications. The fine-grained structure of the explications, on the other hand, were primarily based on the semantic consultation data. These fine-grained meaning nuances and negative evidence about which contexts an expression cannot be used in, were not provided by the corpus data.

5. Results and analysis

5.1. Dictionary entries

According to *Gyldendals Røde Ordbog* (Rød) and *Gyldendals Store Ordbog* (Stor), the German adjective *traurig* can correspond to a broad range of Danish adjectives: *ked af det*, *bedrøvet*, *sørgelig*, *trist*, *bedrøvelig*, *ynkelig* (Rød) and *bedrøvet*, *sørgmodig*, *sørgelig*, *ulykkelig*, *ussel*, and *ynkelig* (Stor). The only examples given for predicative use are *ked af det* and *bedrøvet*, with all others showing attributive use. Semantically, *traurig* can be combined not only with persons, but with abstracts like news, childhood, or circumstances, or body parts (such as eyes), and it appears in several (semi-)idiomatic expressions like *traurige Rolle* “‘sad’ role”, and *trauriges Überbleibsel* “‘sad’ remainder”. For the German adjective *trist* Gyldendal gives the Danish correspondences *trist* and *trøstesløs*—both in predicative position in the examples. Semantically, German *trist* is

used as an attribute for surroundings or a town. The Danish adjective *trist* corresponds to the German adjectives *traurig*, *melancholisch*, *trübe*, *tragisch*, *trist*, *düster*, *freudlos*, *reizlos*, *öde*, *elend*, *miserable*, *ärmlich*, *dürftig*, and *kläglich*. DDO lists also the Danish adjectival expression *ked af det* as a lemma and the only corresponding German words for it are *traurig* and *schweremütig*.

In the Danish monolingual dictionary DDO, the lexical unit *ked af det* is defined as “*ulykkelig eller nedtrykt pga. modgang, skuffelse eller sorg*” ‘unhappy or down because of hardship, disappointment or grief’ and two synonyms, *bedrøvet* and *sørgmodig* are listed. The adjective *trist* is defined as “*præget af dårligt humør eller dårlig stemning*” ‘characterized by bad mood’. The synonyms given for *trist* are *nedtrykt* and *bedrøvet*. The adjective *bedrøvet* is defined as “*ked af det; skuffet om person*” ‘sad; disappointed about a person’. The German monolingual dictionary Duden.de defines the adjectives as follows:

traurig as “*Trauer empfindend, ausdrückend, bekümmert, betrübt; in niedergedrückter Stimmung*” (plus “*Trauer, Kummer, Betrübnis hervorruhend, verursachen, schmerzlich, beklagenswert // kümmerlich, erbärmlich*”)—synonyms given are *bedrückt*, *bekümmert*, *betrübt*, *depressiv*; *unglücklich*;

unglücklich as “*nicht glücklich; traurig und deprimiert; niedergeschlagen*” (plus “*nicht vom Glück begünstigt, ungünstig widrig; ungeschickt*”);

trist as “*durch Öde, Leere, Trostlosigkeit, Eintönigkeit gekennzeichnet; trostlos, freudlos*”, synonyms *abwechslungslos*, *bedrückend*, *deprimierend*, *dunkel*;

betrübt as “*traurig, bekümmert*”, synonyms *bedrückt*, *bekümmert*, *depressiv*, *gedrückt*;

bekümmert as “*traurig, bedrückt, sorgenvoll*”, synonyms *bedrückt*, *betrübt*, *gedrückt*, *sorgenvoll*;

niedergeschlagen as “*durch einen Misserfolg, eine Enttäuschung ratlos, mutlos, traurig*”, synonyms *bedrückt*, *bekümmert*, *betrübt*, *depressiv*;

depressiv as “... *an einer Depression leidend; traurig, niedergeschlagen*” (plus “*durch einen Konjunkturrückgang bestimmt*”); and

deprimiert as “*entmutigt, niedergeschlagen, gedrückt, schwermütig*”, synonyms *depressiv, entmutigt, freudlos, gedrückt*.

Taking the intersection set of this ample lexicographical information, we decided to focus on Danish *ked af det, trist, nedtrykt, deprimeret, depressiv, sørgelig, and ulykkelig*, as well as German *traurig, trist, niedergeschlagen, deprimiert, depressiv, and unglücklich* for further investigation.

5.2. Corpus data and analysis

Table 1 summarizes the **predicative use** of the adjectives in the construction “somebody/something (subject) is ADJ”.

Table 1: *Predicative use of the adjectives (adapted from: http://visl.sdu.dk/semantic_prototypes_overview.pdf)*

- *** prototypical combination
- ** common combination
- * examples can be found
- (*) examples are rare, or based on introspection
- (.) examples could not be found in the corpus but elsewhere on the internet

The semantic classification markers have the following meanings: **act**=action, **act-s**=speech act, **anat**=anatomical (body part), **event**=event, **f-right**=“legal” feature (right or obligation), **feat**=feature, **feat-h**=human feature, **hum**=human person, **Labs**=abstract location, **Lciv**=civitas (town or country), **Lh**=human location, **per**=period, **role**=role, **sick**=disease, **sem**=semantical (semiotic artefact, e.g. readable or listenable), **sem-c**=cognitive semantical (internal), **sem-s**=speech semantical, **soa**=state of affairs, **state-h**=human state.

<i>trist</i>	<i>deprimeret</i>	<i>nedtrykt</i>	<i>sørgelig</i>	<i>depressiv</i>	<i>ked af det</i>	<i>ulykkelig</i>	Dan. subject	sem. class	Ger. subject	<i>traurig</i>	<i>trist</i>	<i>depressiv</i>	<i>deprimiert</i>	<i>niedergeschlagen</i>	<i>betriibt</i>	<i>unglücklich</i>
**	***	***		***	***	**	+HUM/person	hum	+HUM/Person	***		**	**	**	**	**
***			***				<i>det at / det er</i>	clause	<i>es ist / das ist</i>	***						*
								feat-h	<i>Seele</i>						*	
(*)		(*)					<i>Stemming</i>	state-h	<i>Stimmung</i>	(*)	(*)	(*)	(*)	(*)	*	
*							<i>øjne, hjerte, ansigt</i>	anat	<i>(Auge, Herz) Gesicht</i>				(*)	(*)	(*)	(*)
*							<i>verden, livet, tilværelse,</i>	per	<i>(Leben), Gegenwart</i>		*					
*							<i>vejret, sommer</i>	wea	<i>(Wetter) Samstag</i>	(*)	*		*			
*							<i>lufthavne, sceneriet</i>	Lh	<i>Wohngegend</i>		*					
*							<i>nyheder/meningsmåling</i>	sem-c	<i>Entscheidung</i>	(*)						(*)
			**				<i>historie, episode, roman</i>	sem	<i>Kechiche</i>	(*)						(*)
			*				<i>skæbne, manderolle</i>	role	<i>– (Schicksal)</i>	(.)						
			*				<i>Begravelse</i>	event	<i>Beerdigung</i>	*	(*)					
						*	<i>Udelukkelse, hjemsendelse</i>	action	<i>Begriff, Auswahl</i>							*
						*	<i>Udvikling</i>	event	<i>Vorfall, Niederlage</i>	(*)						*
						*	<i>Situation, ægteskab</i>	soa	<i>– (Ehe)</i>							(.)
						*	<i>Kærlighed</i>	f-psych	<i>– (Liebe)</i>							(.)
				*			<i>Marked</i>	Lh	<i>– (Börse)</i>			(*)				
					*		<i>bankerne, regeringen</i>	inst/org								

Table 1: Predicative use of the adjectives (adapted from: http://visl.sdu.dk/semantic_prototypes_overview.pdf)

Danish *trist* is used predicatively—as in *det er trist* ‘this is sad’—with human subjects and clauses as subjects, which is not possible with German *trist*. In these cases, German uses *traurig*. The adjective *trist* is non-literally used in much the same way in both languages, but only Danish uses it for human experience (face, situations).

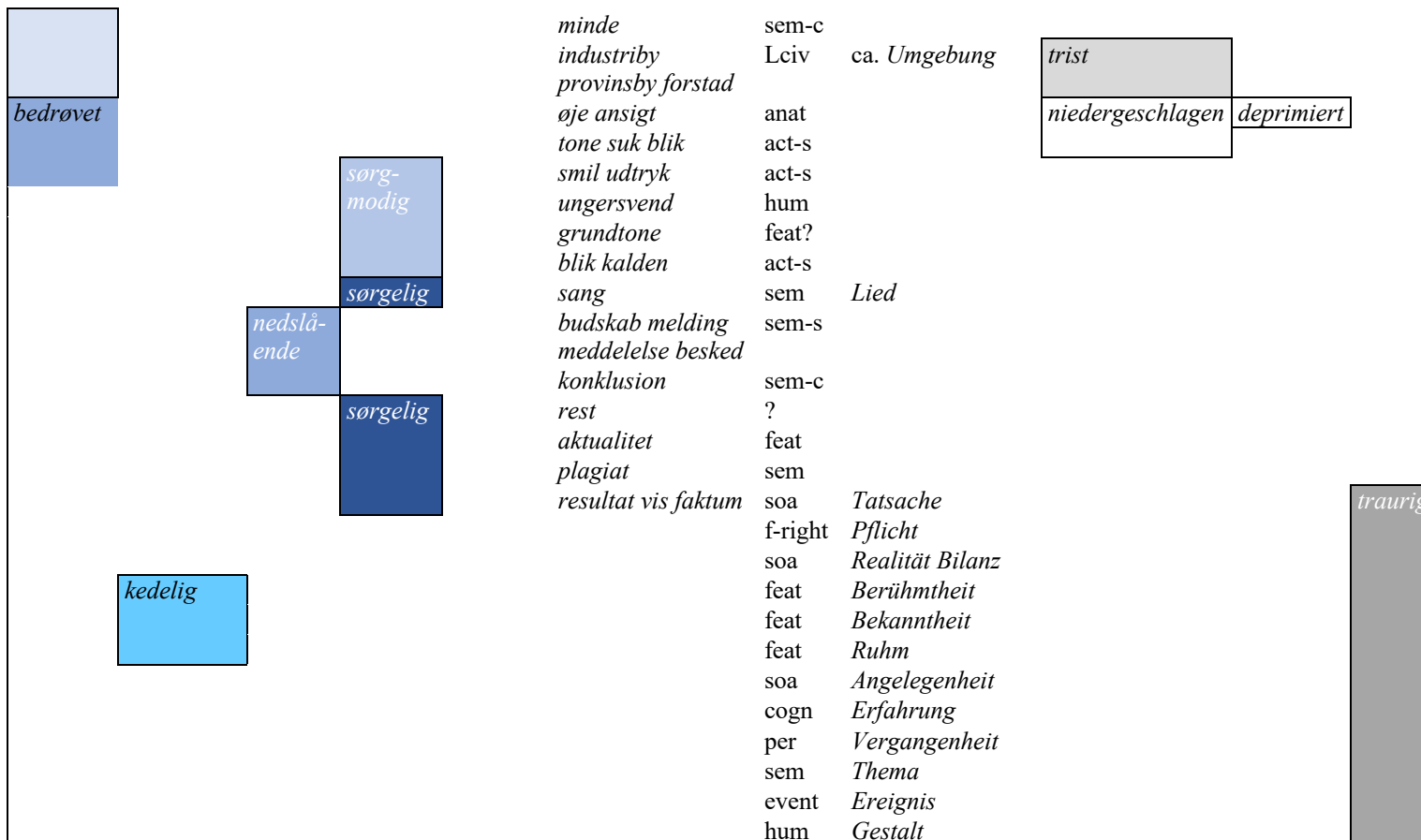
Danish *sørgelig* and German *traurig* have a similar distribution (semiotic products, clauses), but while this is the core meaning of *sørgelig* ‘deplorable’, it is a non-literal meaning for *traurig*, and the latter does not share its literal (human person) connotation with *sørgelig*, matching *ked af det* instead. Still, even this is not a perfect equivalent, since the latter—unlike *traurig*—extends to organizations and institutions.

For Danish *nedtrykt* and German *depressiv*, there are few non-literal uses without a human referent, and for Danish *deprimeret* there are none at all.

Danish *ulykkelig* and German *unglücklich* are used for actions and events, as well as situations and the word *love*. This holds both attributively and predicatively, but for predicatives, non-literal use is limited to transfer of the human sentiment to carriers of that sentiment (face, stories, events), while very few of the predicative examples in either language are used in the sense of “bad”, which is encountered for prenominal use for *trist/sørgelig/traurig*.

Table 2 summarizes the attributive, also called prenominal, use of the adjectives in the construction ADJ + NOUN, where the noun is the head in a nominal phrase and the adjective is the attribute of the head noun.

Dan. adj		Dan. head noun	Sem. class	Ger. head noun	Ger. adj
		<i>sjæl</i>	hum?		
		<i>sindsstemning</i>	state-h	<i>Stimmungslage</i>	<i>deprimiert</i>
		<i>sindstilstand</i>			
		<i>humør</i>			
		<i>tilstand</i>	state (-h?)		
	<i>deprimeret</i>	<i>stemning</i>	state-h	<i>Stimmung</i>	<i>depressiv</i> <i>deprimiert</i> <i>traurig</i>
<i>trist</i>		<i>hum</i>	hum		<i>niedergeschlagen</i>
		<i>hverdag tilværelse</i>	per	<i>Alltag</i>	<i>trist</i>
		<i>dag</i>			
		<i>rekord</i>	Labs	<i>Rekord</i>	
				<i>Höhepunkt</i>	
		<i>skæbne</i>	role	<i>Schicksal</i>	
		<i>historie nyhed</i>	sem	<i>Geschichte</i>	
		<i>kendsgerning</i>		<i>Nachricht</i>	
				<i>Tatsache</i>	
		<i>eksempel sandhed</i>	sem	<i>Beispiel</i>	
		<i>syn</i>	sem-c		
		<i>affære</i>	soa		
		<i>læsning</i>	sem		
		<i>begivenhed</i>	event	<i>Ereignis/Episode</i>	<i>traurig</i>
		<i>oplevelse</i>			
		<i>kapitel</i>	sem	<i>Kapitel</i>	
		<i>konstatering tanke</i>	sem-c	<i>Feststellung</i>	
		<i>exit</i>	event?		



The first part of Table 2 shows that in the attributive position, the most widely used Danish adjective is *trist* and the most widely used German adjective is *traurig*. These two adjectives overlap in many contexts, nevertheless, there are some contexts where Danish *trist* can be used, but not *traurig*, such as “reading material”, “memory”, or “view”. On the other hand, German *traurig* can also be used with several headwords which are not possible with Danish *trist*, such as “famousness”, “experience”, “past”, “topic”, and “event”, as can be seen in the second part of Table 2.

The noun *Stimmung/stemming* seems to be the noun which can stand together with almost all related Danish adjectives and also German *depressive*. There were no instances of *traurige Stimmung* in our corpus, but there can be found examples on the Internet.

As a general rule, predicative use seems to be more conservative and literal for most adjectives: for example, “+ *HUM*/+*act er kedelig*” (‘boring’, both prenominal and predicative) while attributive use can be more creative “*et kedeligt vejr, en kedelig nyhed*” (‘bad’, only prenominal).

5.3. Semantic consultation data and analysis

The informants were asked about the meaning and use of the adjectives with human headwords, independently of predicative or attributive use. According to the Danish informants, the adjectives *trist*, *ked af det*, *bedrøvet*, and *deprimeret* are all used to express some kind of (e.g. mental, emotional, or physical) discomfort. *Trist* and *deprimeret* are usually used about emotional discomfort, such as sorrow or grief, but not physical discomfort. The difference between *deprimeret* and *trist* is that while a *trist* person keeps their sadness-like emotion to themselves, it is more visible when somebody is *deprimeret*. On the other hand, *ked af det* and *bedrøvet* can, according to the informants, both be used about emotional or physical discomfort—for example, when somebody has been hurt. Both *ked af det* and *bedrøvet* can involve a visible reaction, such as crying, but *ked af det* usually involves more tears than *bedrøvet*. An additional difference between *ked af det* and *trist* seems to be that *ked af det* usually refers to a state here and now, while *trist* tends to be more a trait or characteristic of a person.

When asked about the meaning and use of the German sadness adjectives with human headwords, the informants draw the following picture: *traurig* is clearly the adjective with the broadest range of

meaning and use. *Traurig* can express shorter or longer lasting states of sadness, as well as both milder (or just superficial) and stronger, deeper ones. It can designate everything between an emotion that affects a whole person, or a short-lived incident of negative experience in an otherwise happy state of mind. On the other hand, *trist* can hardly be combined with human headwords. In cases where it is, it forces an interpretation of a boring and dull person, causing a *trist* feeling in the observer (typical collocates of *trist* being *Wetter* ‘weather’, *Farben* ‘colours’, and *Umgebung/Gegend* ‘surroundings’). *depressive* is closely associated with a psychiatric diagnosis, whereas *deprimiert*, like *niedergeschlagen/-gedrückt* or *bedrückt/betrübt*, describes a longer-lasting, but temporary, mood that is typically visible by (for example) facial expression (including crying) or a person’s body posture (note the metaphorical quality of *niedergeschlagen/-gedrückt*). Sadness-related adjectives describing a person’s disposition or character would be *melancholisch* or *schwermütig* (which is not necessarily understood as negative). In general, the German sadness-related adjectives do not comprise a component of action, and they are used to describe mental and not physical discomfort. It is worth noting that *traurig*, despite its strong etymological ties to *Trauer* and *trauern*, is hardly associated with the rather specialized sadness-related expression of mourning when a beloved person has passed away.

5.4. NSM explications

Sadness-related emotions are usually caused by “something bad”, thus “something bad happened” is a common element for all sadness-words. Another element of sadness-like emotions is that the experiencer does not feel comfortable with the bad thing that happened, which in NSM terms can be paraphrased by “doesn’t want things like this to happen”. In addition to the sad person not feeling comfortable with what has happened, they also feel that they aren’t capable of doing anything about it. The concept of sadness, which seems to be common for many sadness-related expressions, has been described by Wierzbicka (1999) by the following paraphrase:

- a. I know: something bad happened
- b. I don’t want things like this to happen
- c. I can’t think now: “I will do something because of this”
- d. I know that I can’t do anything

Besides these basic elements of the sadness scenario, different sadness-related expressions in different languages can include additional elements in their meaning. The English adjective *sad* is paraphrased by Wierzbicka (1999) the following way:

[A] X was *sad*

- a. X felt something
- b. Sometimes a person thinks:
- c. “I know something bad happened
- d. I don’t want things like this to happen
- e. I can’t think: I will do something because of it now
- f. I know that I can’t do anything”
- g. when this person thinks this this person feels something bad
- h. X felt something like this

According to the corpus data, in the predicative use, the most relevant expressions that can be combined with the broadest range of nouns are Danish *ked af det* and *trist*, and German *traurig*. The corpus analysis shows that the subjects of these adjectives in predicative constructions can be persons, body parts—such as eyes or heart—or locations or other abstract nouns, such as life, weather, feeling etc.

Common elements of *ked af det* and *trist* used with human subjects are X felt something, X does not want things like this to happen, and because of this, this person feels something bad. Some informants pointed out that *ked af det* was used to express a more or less state-like sadness at the moment while *trist* was used to express a more trait-like feeling lasting for a long time. Another difference between the two adjectives is that while *trist* includes the inability to do something about the sad situation, *ked af det* seems to lack this component.

Based on the semantic template for emotion words (Goddard 2018:ch3), semantic consultation data and the corpus data NSM explications [B]–[D] were designed for the most frequent Danish and German sadness-related expressions which can be used with persons as headwords. Corpus data confirmed that these adjectives are frequently used with human and non-human headwords and have different meanings in these two cases. The fine-grained differences (in bold face below) were revealed by the semantic consultations.

[B] someone X was *ked af det* (at this time) (Danish)

- a. someone X thought like this at this time:
“I know something bad happened
I don’t want things like this to happen”
- b. Because of this, this someone feels something bad at this time
like someone can feel when they think like this

[C] someone X was *trist* (at this time) (Danish)

- a. Someone X thought like this at this time:
“something bad is happening **to me**
I don’t want things like this to happen
It has been like this for **a long time**
I can’t think: I will do something because of it now
I know that I can’t do anything”
- b. Because of this, this someone feels something bad at this time
Like someone can feel when they think like this

[D] someone X was *traurig* (at this time) (German)

- a. someone X thought like this at this time:
“I know something bad happened
I don’t want things like this to happen
I can’t think: I will do something because of it now
I know that I can’t do anything”
- b. Because of this, this someone X feels something bad at this time
like someone can feel **for some time** when they think like this

For the uses of these adjectives NSM-explications [E]–[G] were created if Y is an object:

[E] something Y is *trist* (Danish)

- a. When someone **thinks of** something Y, this someone thinks like this:
“I feel something bad because of Y”
- b. because of this something Y, someone can feel *ked af det*

Danish *trist* can be used together with headwords denoting events or readable or listenable semiotic artefacts which can cause people to feel sad.

[F] something Y is *trist* (German)

- a. When someone **sees** things like Y, this someone thinks like this:
“I feel something bad because of Y
I do not want to be in a place that is like Y”
- b. Because of this something Y this someone thinks like this:
“I do not want to know about Y”

[G] Y is *traurig* (German)

- a. When someone **thinks of** something Y, this someone thinks like this:
“I feel something bad because of Y”
- b. because of this something Y, someone can feel *Traurig*

Explications [E] and [G] are so short because they actually embrace explications [B] and [D].

6. Limitations

We have only made the NSM analysis for the non-metaphorical uses of the most relevant adjectives. In order to get a fuller picture, the metaphorical uses of the same adjectives and some other related adjectives could be analyzed as well. Another limitation of this study is that we have only analyzed adjectives. Onomasiological analyses usually include words from more than one word class. In order to get a more complete picture, it would be necessary to analyze nouns, verbs, and adverbs in addition to adjectives.

7. Discussion

7.1. Differences and similarities between the German and the Danish sadness adjectives

German *traurig* is a general term for sadness, since it can be used with headwords denoting humans, events, states, and readable and listenable semiotic artefacts etc. and it can also be used predicatively with a subordinate clause. Danish, on the contrary, does not have such a general “sadness-term”. Danish, in fact, lacks a monolexemic adjective for describing the sadness-like emotion of humans. The most frequently used sadness-related expression in Danish is the multi-word expression *ked af det*. The use of *ked af det* is syntactically restricted since it only occurs predicatively and not attributively, and it is semantically restricted to persons. The NSM explications have shown that German *traurig* includes passivity and helplessness which is not part of Danish *ked af det*.

Slobin (2006:68) suggested that “it is presumably less demanding to access a single lexical item, such as ‘tiptoe’, than expressions such as ‘on the tips of the toes’, ‘moving quietly and carefully’, etc.” From a cognitive point of view, this means that, compared to German, the concept of sadness-related emotions seems to be less salient in Danish, since it is expressed by a multi-word unit which is more difficult to process than a single word (Slobin 2006). This hypothesis ought to be tested by measuring response time for processing *ked af det* and other one-word expressions for sadness-like emotions.

The meaning of adjectives can also be examined in the light of their antonyms. An antonym of Danish *ked af det* is *lykkelig*, which according to Levisen (2012) means either an intense emotion of a fortunate experiencer, or a life quality of feeling good. Another antonym is *glad*, denoting a carefree mental state when the person does not have to think that bad things can happen, which resembles other Danish cultural keywords *hygge* and *tryk* (Levisen 2012:217). The important Danish cultural value that people can feel good because good things are happening, and they do not have to think about bad things is possibly related to the fact that there is no Danish word for “sad” and the meaning of the multi-word expression *ked af det* does not include helplessness.

7.2. Corpus analysis, semantic consultations and the NSM approach

Wierzbicka (1999:41) suggests that it is more useful to use corpus data for creating NSM explications than to ask informants:

If one asks native speakers of Russian what the difference between grust and pečal is, they usually reply, somewhat vaguely, that one of these emotions is “more concrete”, “more general” than the other, or “more serious”, “more definite”, “more general”, and so on. But a systematic study of the differences in collocations and grammatical frames of the two words and their derivational families allows us to capture the semantic differences in questions in more precise terms.

Wierzbicka (1999), in her study on sadness-related expressions in English and Russian, analyzes some corpus examples but does not actually conduct a systematic quantitative corpus study. A systematic study of the differences in collocations has actually not been used

before, as a basis for writing NSM explications. We made an attempt to combine corpus data about typical syntactic collocates and co-occurrence frequencies with semantic consultation data to create NSM explications for the Danish adjective *trist* and adjective construction *ked af det* and the German adjectives *trist* and *traurig*. We wanted to find the answer to the question of how systematic corpus data analysis can help us in formulating NSM explications. We found that the systematic analysis of typical syntactic collocates can help us in two ways.

Firstly, it helps with sorting out and differentiating between the meanings of a word. If an adjective is used together with human and abstract head nouns, or subjects, if used predicatively, different meanings of the adjective can be identified and disambiguated. This means that the analysis of collocates in the corpus, in line with Goddard's (2000) description of polysemy in the NSM approach, can help us to show how many senses of the same word should be kept apart—that is, how many NSM paraphrases should be written.

Secondly, the systematic semantic analysis of the collocates of the adjectives also helped us to formulate the NSM paraphrases. The NSM paraphrase of the Danish adjective *trist*, when used with a human subject, includes the semantic prime “feel”, but this definition would not cover uses with abstract subjects. When used with an abstract subject in the semantic category “readable or listenable semantic artefacts” such as *news* or *weather*, the NSM paraphrase would include the prime “because of this someone can feel something bad”, because the grammatical subject of the predicative construction denotes an abstract entity which cannot have feelings. Our corpus analysis included the semantic analysis of first order collocates of the adjectives. First order collocates are the possible headwords (nouns) that the adjectives could be used with attributively and the possible subjects that the adjectives could be used with predicatively. In addition to the first order collocates, it is possible to include the broader context of the analyzed words into the analysis and extract more elements of the NSM explication from the corpus. This can be done manually by choosing some particular corpus examples and analyzing them, as seen in Tissari and colleagues (2019), and could possibly be done in a more automated way by taking second order collocates (Heylen et al. 2015) into account, in order to find further contextual cues for the NSM paraphrases. This could be a possible

future question to explore the possibilities of combining vector-based quantitative corpus analysis as a tool for writing NSM explications.

As mentioned by Wierzbicka (1999), besides corpus data, semantic consultations can be used as a basis for creating NSM paraphrases, although semantic consultations can be much like introspection, which is not a reliable data collection method. In order to make semantic consultations more reliable for this purpose, more informants have to be asked more systematically. A further way of using informants to validate the NSM explications is to ask them about the explications previously created by the linguists, as suggested by Gladkova and colleagues (2015).

All in all, we can say that collocation analysis and semantic consultation analysis can both be fruitfully used and supplement each other in creating NSM explications. Collocation analysis shows how many different NSM explications should be written for a lexical unit: for example, *ked af det* can only be combined with human headwords, therefore we only need one explication. German *traurig*, on the other hand, needs two explications. The semantic categorization of the most frequent collocates gives some keys in finding which semantic primes should be included into the explications. When the headword is human, it feels something, while when the headword is non-human, there is a human who feels something because they think about the entity denoted by the non-human headword. Semantic consultations, on the other hand, can help identify more subtle meaning components, which goes beyond collocation analysis. Based on the above experiment, we suggest that different types of systematic corpus analysis and systematic semantic consultations can be successfully combined to create NSM paraphrases for the disambiguation and the definition of word meanings.

References

- Bick, Eckhard. 2009. DeepDict – A Graphical Corpus-based Dictionary of Word Relations. *NODALIDA 2009 Conference Proceedings*, 268–271.
- Dornseiff, Franz. 2012. *Der Deutsche Wortschatz nach Sachgruppen*. Berlin: De Gruyter.

- Ekman, Paul, Wallace V. Friesen and Phoebe Ellsworth. 2013. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. Now York: Pergamon Press.
- Geyer, K. 2015. SMiK-Posterpräsentation von der Konferenz {DEUTSCH} und {DÄNISCH} im Stereotyp: Stereotypenwelten und ihre sprachlich-kulturellen Konstituierungsformen, 25–27. Februar 2015 an der Süddänischen Universität in Odense. In Erla Hallsteinsdóttir and Jörg Kilian (eds.), *SMiK-Projekt*, Syddansk Universitet. Retrieved from <https://www.stereotypenprojekt.eu/app/download/7914051586/SMiK-Posterpublikation.pdf?t=1431025496>
- Gladkova, Anna, Ulla Vanhatalo and Cliff Goddard. 2015. The semantics of interjections: An experimental study with natural semantic metalanguage. *Applied Psycholinguistics* 37(4), 841–65. <https://doi.org/10.1017/S0142716415000260>
- Goddard, Cliff. 2000. Polysemy: a problem of definition. In Yael Ravin and Claudia Leacock (eds.), *Polysemy, theoretical and computational approaches*, 129–51. Oxford: Oxford University Press.
- Goddard, Cliff. 2007. A culture-neutral metalanguage for mental state concepts. In Andrea C. Schalley and Drew Khlentzos (eds.), *Mental states*, 11–35. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.93.04god>
- Goddard, Cliff. 2018. *Ten lectures on the natural semantic metalanguage: Exploring language, thought and culture using simple, translatable words*. Amsterdam: Brill.
- Goddard, Cliff and Susanna Karlsson. 2008. Re-thinking THINK in contrastive perspective: Swedish vs. English. In Cliff Goddard (ed.), *Cross-linguistic semantics*, 225–41. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.102.14god>
- Heylen, Kris, Thomas Wielfaert, Dirk Speelman and Dirk Geeraerts. 2015. Monitoring polysemy: Word space models as a tool for large-scale lexical semantic analysis. *Lingua*, 153–172. <https://doi.org/10.1016/j.lingua.2014.12.001>

- Kemper, Theodore D. 1987. How many emotions are there? Wedding the social and the autonomic components. *American Journal of Sociology* 93(2), 263–89. <https://doi.org/10.1086/228745>
- Levisen, Carsten. 2012. *Cultural semantics and social cognition: A case study on the Danish universe of meaning*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110294651>
- Levisen, Carsten. 2014. The story of Danish happiness: Global discourse and local semantics. *International Journal of Language and Culture*, 174–193. <https://doi.org/10.1075/ijolc.1.2.03lev>
- Levisen, Carsten. 2018. *Pæn, flot, dejlig, and lækker*: A lexical anthropology of Danish folk aesthetics. Paper presented at the VIII International Symposium on Intercultural, Cognitive and Social Pragmatics, 2–4 May, Sevilla, Spain.
- Levisen, Carsten and Sophia Waters. 2017. *Cultural keywords in discourse*. Amsterdam: John Benjamins. <https://doi.org/10.1075/pbns.277>
- Levy, Robert I. 1984. The emotions in comparative perspective. In Klaus R. Scherer and Paul Ekman (eds.), *Approaches to emotion*, 397–412. Hillsdale, NJ: Erlbaum.
- Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S. Corrado and Jeff D. Dean. 2013. Distributed representations of words and phrases and their compositionality. *Neural Information Processing Systems (NIPS)*, 3111–19.
- Nimb, Sanni. (ed.). 2014. *Den Danske Begrebsordbog*. København: Det Danske Sprog- og Litteraturselskab.
- Russell, James A. 1991. In defense of a prototype approach to emotion concepts. *Journal of Personality and Social Psychology* 60, 37–47. <https://doi.org/10.1037/0022-3514.60.1.37>
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Slobin, Dan I. 2006. What makes manner of motion salient? In Maya Hickmann and Stéphane Robert (eds.), *Space in Languages : Linguistic Systems and Cognitive Categories*, 59–81.

Amsterdam: John Benjamins Publishing Company.

<https://doi.org/10.1075/tsl.66.05slo>

Tissari, Heli, Ulla Vanhatalo and Mari Siirainen. 2019. From corpus-assisted to corpus-driven NSM explications: The case of Finnish *viha* (anger, hate). *Lege Artis: Language Yesterday, Today, Tomorrow* 4(1), 290–334.

Wierzbicka, Anna. 1997. *Understanding cultures through their key words: English, Russian, Polish, German and Japanese*. Oxford: Oxford University Press.

Wierzbicka, Anna. 1998. “Sadness” and “anger” in Russian: The non-universality of the so-called “basic human emotions.” In A. Athanasiadou & E. Tabakowska (Eds.), *Speaking of emotions: Conceptualisation and expression* (pp. 3–28). Mouton de Gruyter.

Wierzbicka, Anna. 1999. *Emotions across languages and cultures*. Cambridge: Cambridge University Press.

<https://doi.org/10.1017/CBO9780511521256>

Wierzbicka, Anna. 2009. Overcoming Anglocentrism in emotion research. *Emotion Review* 1(1), 21–23.

<https://doi.org/10.1177/1754073908097179>

Ye, Zhengdao. 2013. Comparing the natural semantic metalanguage (NSM) approach to emotion and the GRID paradigm. In Johnny J.R. Fontaine, Klaus R. Scherer and Cristina Soriano (eds.), *Components of emotional meaning: A sourcebook*, 339–409. Oxford: Oxford University Press.

<https://doi.org/10.1093/acprof:oso/9780199592746.003.0028>

Zhang, Ruihua. 2016. *Sadness expressions in English and Chinese: Corpus linguistic contrastive semantic analysis*. New York: Bloomsbury Publishing.