

Igor A. Mel'chuk
Université de Montréal

Leo Wanner
Universität Stuttgart

Towards an Efficient Representation of Restricted Lexical Cooccurrence¹

Abstract

In this paper, we address the problem of efficient representation of restricted lexical cooccurrence information in the lexicon. In a study carried out with forty German emotion lexemes, significant correlations between the semantics of these lexemes and their verbal collocates have been found. Drawing on these correlations, we suggest a more efficient representation of collocation information in the lexicon: recurrent collocates should be extracted from the individual lexical entries and listed together with their semantic constraints in the entry for what we call the generic lexeme of the semantic field under study.

1. Introduction

This paper addresses the problem of efficient representation of restricted lexical cooccurrence information in dictionaries. Restricted lexical cooccurrence (or collocation, cf. Hausmann 1985; Benson 1989) is the cooccurrence of lexemes such that the choice of a specific lexeme L_1 for the expression of a given meaning is contingent on another lexeme L_2 to which this meaning is applied. Thus, e.g., the choice of *fierce* for the expression of an 'intense' resistance is contingent on *resistance* (cf. *fierce resistance*). In a collocation $L_1 + L_2$, the L_1 is called 'the collocate' (in *fierce resistance*, this is *fierce*) and L_2 – 'the base' (in *fierce resistance*, this is *resistance*).

In accordance with the arbitrariness of collocations (see, e.g., Mel'chuk 1988; Hausmann 1985) they cannot, generally speaking, be automatically predicted from the meaning or grammatical properties of the base. However, this arbitrariness does not preclude partial motivation. Thus, collocates may correlate with the semantic class of their base. For instance, in French, most nominal lexemes that denote emotional attitudes go with *éprouver* '[to] feel' (*éprouver un profond respect/mépris* '[to] feel deep respect/contempt', *éprouver de la compassion* '[to] feel compassion', etc.). In English, *[to] issue* combines not only with REPORT but also with STATEMENT, COMMENT, ORDER, etc.; *[to] reject* goes with PLEA as well as with PROPOSAL, ADVICE, SUPPOSITION, OFFER, and so on.² This phenomenon is quite frequent in language; therefore, it would be desirable to generalize restricted lexical cooccurrence along semantic lines.

In this paper, we present a study carried out on emotion lexemes in German. In this study, significant correlations between the semantics of forty emotion nouns and their cooccurrence with twenty-five verbal lexemes have been found. These correlations have been used to extract recurrent collocates from the individual lexical entries in order to list them in the entry for what is called the generic lexeme of the semantic field under study. Such treatment of lexical data allows for a reasonable generalization of collocation information. At the same time the study has shown that the correlations are far from being universal: idiosyncrasies abound in collocations, so that many of them have to be simply listed.

Due to space limits, we give in what follows only a few examples and a very short theoretical outline of our work; for a comprehensive discussion, see (Mel'chuk & Wanner, to appear).

2. Representing lexical cooccurrence in the ECD

The theoretical framework of our study is the *Meaning-Text Theory* (MTT) (Mel'chuk 1981) and, more precisely, the lexicographic part of MTT – the *Explanatory Combinatorial Dictionary* (ECD) (cf., e.g., Mel'chuk & Polguère 1987), which offers a systematic and thorough description of collocation phenomena.

It is impossible to state in this paper the general postulates, main principles and the formal apparatus advanced in MTT and ECD so we will rely on the references given in the course of this article; here we will only say a few words about the representation of collocations in the *Lexical Cooccurrence Zone* of an ECD entry under the heading of *Lexical Functions* (LFS), cf. (Mel'chuk, forthcoming).

In terms of LFS, the collocation L_1 (=collocate) + L_2 (=base) is presented as $f(L_2) = L_1$, where L_2 is called the **keyword** of the corresponding LF and L_1 its value. An LF represents a very general semantico-syntactic relation between L_1 and L_2 . About sixty LFS which represent standard semantico-syntactic relations have been found across different languages.

Two examples shall illustrate the concept of LFS:

1. **Oper₁**: provides for its keyword L_2 (which is a predicate noun, i.e. a noun denoting an action, an event, a state, etc.) a verb L_1 with the meaning 'perform', 'undergo', 'be in a state', etc. L_1 takes as its grammatical subject the name of the first actant of L, i.e. the agent of the action, the undergoer of the event, the subject of the state, etc.; and as its direct object, the lexeme L_2 itself.

Examples of Oper₁ for nouns meaning 'complaint' are:

Eng.	Oper ₁ (<i>complaint</i>)	= [to] lodge [DET ~]
Fr.	Oper ₁ (<i>plainte</i>)	= porter [~] ' [to] carry '
Germ.	Oper ₁ (<i>Beschwerde</i>)	= [DET ~ _{acc}] einreichen ' [to] hand in '
Rus.	Oper ₁ (<i>žaloba</i>)	= podat [~u] ' [to] hand in '

The expression in square brackets is the syntactic Government Pattern of the LF verb in question; 'N' replaces the keyword; and DET indicates that the noun takes the determination according to general rules of the language.

2. **FinFunc₀** provides for its keyword L₂ (which is again a predicate noun) a verb L₁ with the meaning 'cease to take place'. L₁ accepts as its grammatical subject L₂ itself.

Examples of FinFunc₀ for nouns denoting 'fear' are:

Engl.	FinFunc ₀ (<i>fear</i>)	[to] wear off, subside
Fr.	FinFunc ₀ (<i>peur</i>)	se calmer ' [to] calm down '
Germ.	FinFunc ₀ (<i>Angst</i>)	sich legen ' [to] lie down '
It.	FinFunc ₀ (<i>paura</i>)	scomparire ' [to] disappear '

3. Emotion lexemes in German

3.1 The notion of semantic dimension

In psychology, the meaning of emotions and emotion lexemes has often been described via semantic **multidimensional scaling** (cf., for example, Dahl & Stengel 1978; Russel 1980).

A **semantic dimension** is a set of two or three mutually exclusive values; all but one of these values are "marked". A marked value of a dimension **n** is a label which stands for a semantic component of the corresponding definition of a lexeme. The unmarked value of **n** means that this dimension is irrelevant to the definition in question; we will call it "n-neutral". To describe the semantics of German emotion lexemes for the purpose of our study, i.e. for establishing correlations between meaning and restricted lexical cooccurrence, we propose eleven semantic dimensions with the following values:³

1 INTENSITY	= { 'intense', 'moderate', 'intensity-neutral' }
2 POLARITY	= { 'pleasant', 'unpleasant', 'polarity-neutral' }
3 MANIFESTABILITY	= { 'manifested', 'manifestable', 'manifestation-neutral' }
4 DIRECTIONALITY	= ('directed at', 'direction-neutral')
5 MENTALITY	= { 'mental', 'mentality-neutral' }
6 REACTIVITY	= { 'reactive', 'reactivity-neutral' }
7 ATTITUDINALITY	= { 'attitudinal', 'attitude-neutral' }

8 ACTIVITY	= {'active', 'activity-neutral'}
9 EXCITATION	= {'excited-state', 'excitation-neutral'}
10 SELF-CONTROL	= {'self-control-loss-inflicting', 'self-control-neutral'}
11 PERMANENCE	= {'temporary', 'permanent'}

Examples of 'intense' emotion lexemes include: BEGEISTERUNG 'enthusiasm', EMPÖRUNG 'indignation', ENTSETZEN 'horror'; 'moderate' are the following emotion lexemes: VERÄRGERUNG 'annoyance', VERWUNDERUNG 'amazement', and VERDRUSS 'vexation'; and 'intensity-neutral' are, e.g., ACHTUNG 'respect', ANGST 'fear', NEID 'envy', etc.

There are substantial implications between the dimensions listed above. For instance, an 'attitudinal' emotion lexeme (such as ACHTUNG 'respect') is necessarily 'directed at' and 'permanent'. However, we have not studied these implications as yet, and, therefore, do not consider them systematically in what follows.

The eleven dimensions have been arrived at by analyzing in parallel the definitions of German emotion lexemes, which are written in the ECD-style (see, e.g., Mel'chuk et al., 1984, 1988, 1992) taking findings in lexical semantics, lexicography and psychology into account, and the values of their LFS. For reasons of typographical convenience, we represent them as linear sequences of explicit conventional labels in which all n-neutral values are omitted. Such a sequence is the "definitorial" part of what we call an **abridged definition** of a lexeme. For example, the abridged definition of ACHTUNG 'respect' appears as follows:

Achtung
 von X vor Y 'X's respect towards Y' = X's pleasant, mental, attitudinal,
 active, permanent emotion directed at Y

Let it be emphasized that an abridged definition is, strictly speaking, redundant with respect to the full definition as given in an ECD, see, e.g., (Mel'chuk & Polguère 1987). But it presents the information necessary for our specific task in an explicit, standard and compact form, which significantly facilitates the finding of correlations between lexical cooccurrence and meaning.

3.2 Restricted lexical cooccurrence of emotion lexemes in German

The cooccurrence data used in our study have partially been obtained from the corpora of the Institut für Deutsche Sprache (IDS), Mannheim; in cases where these corpora did not give sufficient evidence for the acceptability of specific collocations, the judgements have been left to the linguistic intuition of the authors. Their decisions as to what is possible/impossible in this domain were checked by ten native speakers and buttressed by a few findings from literary texts.

In the following, we briefly consider three (out of twenty-five) collocates of our emotion lexemes: *empfinden* '[to] perceive', *entgegenbringen* '[to] show', and *schüren* '[to] fan'.

Empfinden '[to] perceive' (= $\text{Oper}_1(L_2)$) is one of the most current choices among other Verb + Noun expressions for emotion lexemes: it combines with thirty-five out of the forty lexemes under consideration. The lexemes with which it does not co-occur are: *AUFREGUNG* 'excitement', *ENTZÜCKEN* 'delight', *PANIK* 'panic', *STAUNEN* 'astonishment', and *VERWUNDERUNG* 'amazement'; these lexemes do not form a semantic subclass.

Entgegenbringen '[to] show' (= $\text{Oper}_1(L_2)$) cooccurs with 'attitudinal' emotion lexemes. These are: *ACHTUNG* 'respect', *HASS* 'hatred', *LIEBE* 'love', *MITLEID* 'compassion', *VERACHTUNG* 'contempt', and *ZUNEIGUNG* 'affection'.

Schüren '[to] fan' (= $\text{CausContFunc}_1(L_2)$) expresses, e.g., a disapproval of the emotion in question by the speaker. Therefore, it is natural for it to co-occur mostly with 'unpleasant' emotion lexemes (e.g., *ANGST* 'fear', *EIFERSUCHT* 'jealousy', *GROLL* 'grudge', etc.). But nevertheless the subset of 'unpleasant' lexemes that actually co-occur with *schüren* is considerably smaller than the subset of those which do not (11:20). Additionally, *schüren* can also co-occur with 'pleasant' emotion lexemes (such as, e.g., *LEIDENSCHAFT* 'passion': *eine krankhafte Leidenschaft schüren* '[to] fan a sick passion', where *schüren* implies that the speaker strongly disapproves of the passion although it could be very pleasant for the Experiencer). To summarize: *schüren* shows an extremely heterogeneous behavior.

4. Discussion of lexical cooccurrence/meaning correlations

At present, an ECD describes restricted lexical cooccurrence by specifying, for each head lexeme *L*, all the values of all applicable LFS in *L*'s entry explicitly – ignoring possible cooccurrence/meaning correlations and regardless of resulting redundancy. Yet many different emotion lexemes have the same value for a given LF: cf. $\text{Oper}_1(X) = \textit{hegen}$ '[to] harbor' with $X = \{\textit{ACHTUNG}$ 'respect', *GROLL* 'grudge', *HOFFNUNG* 'hope', *LEIDENSCHAFT* 'passion', *ZUNEIGUNG* 'affection'\}; see also *entgegenbringen* above.

This gives rise to the two following questions:

1. Is it possible to avoid the tedious repetition of the same value of a given LF *f* by specifying this value just once for many or even for all of the keywords concerned, thus gaining a substantive generalization?
2. If yes, how to present in an ECD the generalized specification of LFS?

German emotion lexemes taken as bases behave similarly with respect to at least two collocates: *empfinden* '[to] perceive' and *fühlen* '[to] feel'. Nearly all of them co-occur with *empfinden* (see above), and so does the noun **GEFÜHL** 'emotion'; *fühlen* '[to] feel' is also nearly universal, although less so, for the same lexemes. Therefore, a generalization along the following lines seems possible:

The definitions of all German emotion lexemes contain a common component: 'emotion' expressed by the German lexeme **GEFÜHL**. Thus, **GEFÜHL** is the **generic** lexeme of the semantic field of emotions in German. Then, for emotion lexemes, the verbs *empfinden* and *fühlen* are specified only once – in the lexical entry of the generic lexeme **GEFÜHL** 'emotion' as values of the LF Oper₁. All specific lexemes denoting emotions, i.e. containing the component 'Gefühl' in their definitions, should then **inherit** this value of Oper₁ from the lexeme **GEFÜHL**.

Unfortunately, as a general rule, there is no unique correlation between the values of LFs applicable to the generic lexeme and the values of the same LFs applied to the specific descendants of the latter (cf. Heid & Raab 1989). Thus, even *empfinden* does not combine with five out of forty emotion lexemes and *fühlen* does not with ten (see Mel'chuk & Wanner, to appear). Furthermore, the verb *erleben* goes with **ENTTÄUSCHUNG** 'disappointment', while *schöpfen* and *machen* go with **HOFFNUNG** 'hope'; but none of these verbs combines with **GEFÜHL**:

Oper ₁ (GEFÜHL)	=	<i>empfinden</i> , * <i>erleben</i>
IncepOper ₁ (GEFÜHL)	=	* <i>schöpfen</i>
CausFunc ₁ (GEFÜHL)	=	<i>wecken</i> , * <i>machen</i> , <i>wachrufen</i>

Similar examples can be multiplied endlessly.

The absence of a strict enough correspondence between LF values in the entry for **GEFÜHL** and those in the individual entries for emotion lexemes is due to a very high degree of idiosyncrasy in the LF Verb + Noun collocations.

As a result, we face two extremes: either a whole semantic class of lexemes showing a (nearly) identical restricted lexical cooccurrence (*empfinden* and *fühlen* with **GEFÜHL** lexemes, *entgegenbringen* with 'attitudinal' lexemes); or individual lexemes featuring completely idiosyncratic, non-generalizable cooccurrence (*machen* with **ANGST** 'fear', **FREUDE** 'joy', and **HOFFNUNG** 'hope'). However, in natural language intermediate cases abound: some members of a semantic class show somewhat similar restricted lexical cooccurrence. For instance, **ANGST** 'fear', **BEGEISTERUNG** 'enthusiasm', **ENTTÄUSCHUNG** 'disappointment' co-occur with *erfassen* '[to] overcome' as

IncepFunc₁ – just as many other emotion lexemes do; yet AUFREGUNG ‘excitement’, ENTZÜCKEN ‘delight’, SCHRECK ‘terror’ and a few more do not.

It is such intermediate cases that we deal with in the present paper. In other words, we set out to find some regularities in a domain which is irregular by definition.

To do so, we take two sets of “processed” data:

1. A set of the abridged lexicographic definitions of forty emotion lexemes, i.e. the characterization of the lexemes in terms of eleven semantic “dimensions”.
2. A set of the LF values specifications for the same forty lexemes – restricted to the twenty-five collocate verbs investigated.

Starting from this data, we try to find the optimal correlation between the values of LFs and semantic features in the abridged definitions. (By “optimal” we mean a correlation that ensures maximal generalization with a minimal number of individual exceptions, i.e. the best information compression possible.) As pointed out above, several such correlations exist, so that Question 1 at the beginning of this section has to be answered in the positive: substantive generalizations over LF values can and should be stated. We have now to answer Question 2, i.e. to propose a lexicographic format that is able to cope with such generalizations.

5. Implementing lexical inheritance in an ECD

5.1 Proposal

Our proposal can be stated in terms of the following five steps:

1. Delimiting the semantic field under analysis and preparing full-fledged lexical entries using the ECD methodology for each of the chosen lexemes.
2. Determining the generic lexeme of the field by analyzing the full-fledged definitions and elaborating its lexical entry. This process results in an improvement of the definitions themselves, as it leads to greater standardization.
3. Introducing semantic dimensions capable of capturing relevant semantic commonalities in full-fledged definitions; deriving abridged definitions expressed in terms of semantic dimensions. This step may bring about a further standardization of the original definitions.
4. Extracting semantically motivated commonalities in LF values found in the lexical entries for specific lexemes and transferring them to the lexical entry for the generic lexeme. While doing this, the researcher has to make sure that each transferred element is supplied with semantic conditions which license its use with the specific lexeme it has been extracted from. These conditions are formulated in terms of the semantic dimensions.

The LF values listed in the entry for the generic lexeme are then **inherited** by all lexemes of the field under the control of semantic dimensions. At the same time, all the exceptions – i.e. the common elements of the LF values that cannot be extracted – are explicitly listed in the individual entries.

5. Reorganizing the lexical entry for the generic lexeme by dividing it into two parts: its own lexical entry (describing its own syntax and cooccurrence – the “private” subentry) and the subentry for the extracted commonalities of the field (the “public” subentry, see below).

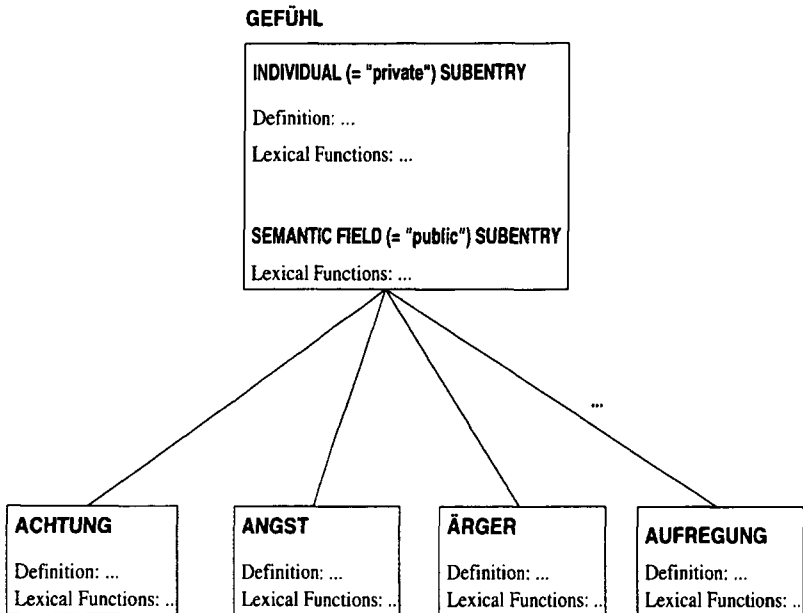


Figure 1: The structure of the lexicon within the field of emotion lexemes

Figure 1 shows the resulting structure of the lexicon within the semantic field of emotion lexemes.

The procedure as described above results in the following treatment of the three collocates discussed in Section 3.2:

Empfinden ‘[to] perceive’ is specified, among other verbs, as an element of the value of Oper₁ in the “public” subentry of the entry for

GEFÜHL ‘emotion’ (the generic lexeme of our semantic field):

Oper₁ : empfinden, fühlen, ...

The exceptions, i.e. the lexemes that do not co-occur with *empfinden* (such as AUFREGUNG ‘excitement’, ENTZÜCKEN ‘delight’, etc.), are marked as such in their own lexical entries via negation (the sign ‘-|’ stands for negation):

Oper₁ : -| empfinden

Entgegenbringen ‘[to] show’ is specified as an element of the value of Oper₁ in the “public” subentry for GEFÜHL ‘emotion’ with the semantic constraint attitudinal, which means that only ‘attitudinal’ emotion lexemes co-occur with it (the ‘|’ sign stands for “condition”):

Oper₁ : entgegenbringen
[DET N_{dat-acc}] | ‘attitudinal’

Schüren ‘[to] fan’ as an element of the value of CausContFunc₁ shows a quite heterogeneous cooccurrence (see above). Unable to find a plausible semantic explanation for this heterogeneous behavior, we specify this cooccurrence explicitly in the entries for the eleven lexemes in question.

5.2 Lexical entries

In this section we give first the entry for the generic lexeme of the semantic field of German emotion lexemes (= GEFÜHL ‘emotion’) and, second, the lexical entry for ANGST ‘fear’ in its “full” and “compressed” forms. Note that these entries consist merely of abridged definitions and the information on lexical cooccurrence. For the sake of simplicity, we have omitted, e.g., syntactic information, that has also been subject to generalization in our work described in (Mel’chuk & Wanner, to appear).

In the individual entries, the sign ‘ ’ (cf. Oper₁) means that the value of the LF in question that is specified locally is to be added to the set of the elements of the inherited LF value. In the absence of the ‘ ’ sign in the entry, the locally specified value of an LF replaces the inherited value of this LF.

The sign ‘-|’ with X being an element of the value of an LF means (as stated above) that this element does not co-occur with the lexeme in question.

N_{acc}, N_{dat}, and N_{gen} stand for nouns in Accusative, Dative, and Genitive, respectively.

5.2.1 The lexical entry for GEFÜHL – the Generic Lexeme of the Semantic Field of Emotions

GEFÜHL₁, neutr

Individual (= “Private”) Subentry

X's Gefühl des W gegenüber Y wegen Z 'X's emotion of W towards Y because of Z' = State W of X's psyche caused by and directed at Y (and causing that X tends to interact with Y because of Z).

Lexical Functions

IncepPredMinus	: nachlassen
Oper ₁	: empfinden, haben [DET ~ _{acc}]
IncepOper ₁	: bekommen [DET ~ _{acc}]
FinFunc ₀	: sich legen
Liqu ₁ Func ₀	: überwinden [PRON _{poss} /DET ~ _{acc}]
IncepFunc ₁	: aufkommen [in N _{dat}]
Magn + IncepFact ₁	: überkommen [N _{acc}]

Semantic Field (= “Public”) Subentry

Lexical Functions

IncepPredMinus	: nachlassen	'excited-state'
Oper ₁	: empfinden, fühlen [DET/Ø ~ _{acc}]; entgegenbringen [N _{dat} DET ~ _{acc}]	'attitudinal', 'permanent'
Magn + IncepOper ₁	: geraten [in ~ _{acc}] ausbrechen [in ~ _{acc}]	'manifested', 'intense' ∧ 'manifested'
FinFunc ₀	: sich legen	'excited-state'
fast FinFunc ₀	: verfliegen	'excited-state'
Liqu ₁ Func ₀	: überwinden [PRON _{poss} /DET ~ _{acc}]	¬'moderate'
IncepFunc ₁	: aufkommen [in N _{dat}]	¬'moderate'
Magn + IncepFunc ₁	: erfassen [N _{acc}]	'self-control-loss-inflicting'
Magn + fast IncepFunc ₁	: packen [N _{acc}]	'mental' ∧ 'reactive' ∧ 'temporary'
CausFunc ₁	: wecken [in N _{dat} ~ _{acc}]	'mental' ∧ 'reactive'
Caus ₂ Func ₁	: hervorrufen [bei N _{dat} ~ _{acc}], erregen [in N _{dat} ~ _{acc}]	
Liqu ₁ Fact ₀	: unterdrücken [PRON _{poss} /DET ~ _{acc}]	
(Magn +) IncepFact ₁	: überkommen [N _{acc}]	

5.2.2 The lexical entry for ANGST 'fear'

In what follows, we present first the lexical entry for ANGST 'fear' with all LFs which are applicable to ANGST; and then, the "compressed" entry from which all generalizable LF values are extracted. For illustration purposes, these LF values are printed in bold face in the first entry.

ANGST, fem (full entry)

Definition

Angst von X vor Y wegen Z = 'X's fear of Y because of Z' = X's unpleasant, reactive, active, excited-state, self-control-loss-inflicting, permanent or temporary Gefühl directed at Y because of Z

Lexical Functions

IncepPredMinus	: nachlassen
Oper ₁	: empfinden, fühlen, haben [\sim_{acc}]
IncepOper ₁	: bekommen [\sim_{acc}]
Caus ₂ Oper ₁	: versetzen [N_{acc} in \sim_{acc}]
FinFunc ₀	: sich legen
fast FinFunc ₀	: verfliegen
Liqu ₁ Func ₀	: überwinden [PRON _{poss} / DET \sim_{acc}]
IncepFunc ₁	: aufkommen [in N_{dat}]
Magn + IncepFunc ₁	: erfassen [N_{acc}]
Magn + fast IncepFunc ₁	: packen [N_{acc}]
CausContFunc ₁	: schüren [in N_{dat} \sim_{acc}]
Caus ₂ Func ₁	: einflößen, erregen [N_{dat} \sim_{acc}], wecken [in N_{dat} \sim_{acc}]
Caus ₍₂₎ Func ₁	: hervorrufen [bei N_{dat} \sim_{acc}], machen [N_{dat} \sim_{acc}]
Liqu ₁ Fact ₀	: unterdrücken [PRON _{poss} / DET \sim_{acc}]
Magn + IncepFact ₁	: überkommen [$\sim N_{acc}$]

ANGST, fem (compressed entry)

Angst von X vor Y wegen Z 'X's fear of Y because of Z' = X's unpleasant, manifestable, reactive, active, excited-state, self-control-loss-inflicting, permanent or temporary Gefühl directed at Y because of Z

IncepOper ₁	: bekommen [\sim_{acc}]
Caus ₂ Oper ₁	: versetzen [N_{acc} in \sim_{acc}]
CausContFunc ₁	: schüren [in N_{dat} \sim_{acc}]
†Caus ₂ Func ₁	: einflößen [N_{dat} \sim_{acc}], wecken [in N_{dat} \sim_{acc}]
Caus ₍₂₎ Func ₁	: machen [N_{dat} \sim_{acc}]

6. Conclusion

Our study has shown that at least in the field of German emotion lexemes generalizations over LF values along semantic lines are possible. These

generalizations are extremely useful since they help to avoid redundancy: once specified in the "public" subentry of the lexical entry for the generic lexeme of a given semantic field, shared LF values need no longer be repeated in individual entries.

However, it has also been shown that it is often impossible to find correlations between lexical cooccurrence of the key lexemes and their semantic features. After all, language is notoriously capricious and unpredictable.

Notes

- 1 This work has partially been supported by Canada Research Council Grant 410-91-1844 (I. Mel'chuk) and the Ministry for Research and Science, Baden-Württemberg (L. Wanner). Many thanks are due to Ulrich Heid for valuable comments on an earlier draft of this paper. We are also very grateful to Monika Lacher for proofreading the final version.
- 2 A substantive dependency between semantic subclasses of lexemes and their collocates has been first stated, to our knowledge, in (Heid & Raab 1989: 132-133), based on a description of French attitudinal emotion lexemes in (Mel'chuk et al., 1984); but see also (Reuther, forthcoming). For a general discussion of correlations between semantics and lexical cooccurrence, see also (Pustejovsky et al., 1993).
- 3 Our semantic description of emotion lexemes draws heavily on research in lexical semantics and lexicography (see, in particular, Iordanskaja 1972, 1973; Wierzbicka 1972, 1992; Mel'chuk et al., 1984; Iordanskaja & Mel'chuk, 1991; Apresjan & Apresjan 1993; and Bergenholtz, 1980). Psychological studies dealing primarily with a cognitive model of emotions (see, e.g., Averill 1975; Ortony et al. 1988; Oatley & Johnson-Laird 1987) have been another important source of inspiration for us.

References

- Apresjan, V.Ju. & Apresjan, Ju. D. 1993. "Metafora v semanticeskom predstavleniiemocij" [Metaphor in the semantic representation of emotions]. In *Voprosy Jazykoznanija*, (3), 27-35.
- Averill, J.R. 1975. "A Semantic Atlas of Emotion Concepts". In *Catalog of Selected Documents in Psychology*, 330(5).
- Benson, M. 1989. "The Structure of the Collocational Dictionary". In *International Journal of Lexicography*, 2(1), 1-14.
- Bergenholtz, H. 1980. *Das Wortfeld "Angst"*. Stuttgart: Klett-Cotta.
- Dahl, H. & Stengel, B. 1978. "A Classification of Emotion Words". In *Psychoanalysis and Contemporary Thought*, 1(2), 269-312.
- Hausmann, F. J. 1985. "Kollokationen im deutschen Wörterbuch: Ein Beitrag zur Theorie des lexikographischen Beispiels". In H. Bergenholtz & J. Mugdan (eds). *Lexikographie und Grammatik, Akten des Essener Kolloquiums zur Grammatik im Wörterbuch*, 118-129. Essen.
- Heid, U. & Raab, S. 1989. "Collocations in Multilingual Generation". In *Proceedings of the 4th. Conference of the European Chapter of the Association for Computational Linguistics*, 130-136. Manchester: Association for Computational Linguistics.
- Iordanskaja, L.N. 1972. Leksikograficeskoe opisanie russkix vyrazenij, oboznacajuscix fiziceskie symptomy cuvstv [Lexicographic description of Russian expressions denoting physical symptoms of emotions]. In *Masinnij perevod i prikladnaja lingvistika*, 16, 3-30. Moscow.
- Iordanskaja, L.N. 1973. Tentative Lexicographic Definitions for a Group of Russian Words Denoting Emotions. In F. Kiefer (ed). *Trends in Soviet Theoretical Linguistics*, 389-410. Dordrecht: Reidel.
- Iordanskaja, L.N. & Mel'chuk, I.A. 1991. Semantics of Two Emotion Verbs in Russian: bojat'sja 'to be afraid' and nadejat'sja 'to hope'. In *Australian Journal of Linguistics*, 10, 307-357.

- Mel'chuk, I.A. forthcoming. Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon. In L. Wanner (ed). *Lexical Functions in Lexicography and Natural Language Processing*.
- Mel'chuk, I.A. 1981. Meaning–text Models: a Recent Trend in Soviet Linguistics. In *Annual Review of Anthropology*, 10, 27–62.
- Mel'chuk, I.A. & Polguère, A. 1987. A Formal Lexicon in the Meaning–Text Theory (or How to Do Lexica with Words). In *Computational Linguistics*, 13(3–4), 276–289.
- Mel'chuk, I.A. & Wanner, L. to appear. Lexical Cooccurrence and Lexical Inheritance. Emotion Lexemes in German: A Lexicographic Case Study. In *Lexikos*, (4), Cape Town.
- Mel'chuk, I.A. *et al.* 1984, 1988, 1992. *Dictionnaire explicatif et combinatoire du français contemporain, I, II, III* Montréal: Presses de l'Université de Montréal.
- Oatley, K. & Johnson–Laird, P.N. 1987. Towards a Cognitive Theory of Emotions. In *Emotion and Cognition*, 1(1), 29–50.
- Ortony, A., Glöer, G. L. & Collins, A. 1988. *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.
- Pustejovsky, J., Bergler, S. & Anick, P. 1993. Lexical Semantic Techniques for Corpus Analysis. In *Computational Linguistics*, 19(2), 331–358.
- Reuther, T. forthcoming. On Dictionary Entries for Support Verbs: The Cases of Russian 'vesti', 'provodit', and 'proizvodit'. In L. Wanner (ed). *Lexical Functions in Lexicography and Natural Language Processing*.
- Russel, J. A. 1980. A Circumplex Model of Affect. In *Journal of Personality and Social Psychology*, 39(6), 1161–1178.
- Wierzbicka, A. 1972. *Semantic Primitives*. Frankfurt am Main: Athenäum Verlag.
- Wierzbicka, A. 1992. Defining Emotion Concepts. In *Cognitive Science*, 16(4), 539–581.