John Benjamins Publishing Company

A two-dimensional approach to the study of motivation in lexical typology and its first application to French high-frequency vocabulary*

Peter Koch and Daniela Marzo
University of Tübingen

'Morphological' and 'semantic' motivation are not just two *types* (Ullmann 1966), but *two* interrelated *dimensions* of lexical motivation. For instance, Fr. *poire* 'pear' — *poirier* 'pear-tree' expresses the same cognitive relation as the polysemy of Russ. *gruša*, and, at the same time, polysemy is only one formal device among others expressing cognitive relations that underlie lexical motivation. So the two dimensions of formal and cognitive relations in motivation only exist in combination. A sub-dimension of the formal aspect of motivation is the degree of formal transparency (cf. Fr. *jouer* 'to play a game' — *jeu* 'game').

This factorization in different dimensions leads to a universally applicable grid for the description of lexical motivation. As a first step of a comparative research project of lexical motivation in different languages it is applied to the 500 most frequent lexical words of French and yields a systematic *motivational profile* of French high-frequency vocabulary. In Section 5.3 this pilot study is discussed in view of an approach to lexical typology which could be applied to any other language.

0.   Introduction: Where do words come from?

In his dialogue *Cratylos* Plato discusses a question that had already been discussed by some of his philosophical precursors: are names "right" in that their formal properties correspond by nature (φύσει) to what they express, or are they purely conventional (συνθήκῃ / ὁμολογίᾳ)? Even though Plato does not definitively make up his mind on this issue, he sketches at least two forms of what we would nowadays call 'motivation' as opposed to 'arbitrariness'.

Complex words (or 'later' words) can be decomposed into their elements ('earliest' words), which, in turn, if they are not thought to be purely conventional, may imitate by voice the "thing" they denominate (Plato 1996: 8–11 = 384C–D, 126–127 = 421A, 132–137 = 423A–424A).

The problem of arbitrariness — in the various shapes it may take — has been a long runner in philosophical and semiotic reflection from antiquity to modern times (Coseriu 2004: 1–62, 99–162). In modern linguistics, Ferdinand de Saussure took up this question within his semiotic framework of the *Cours de linguistique générale* (1916: 100–102, 180–184). Starting from the hypothesis that the linguistic sign is essentially arbitrary (exemplified by Fr. *sœur* 'sister'), he nevertheless admits that this fundamental arbitrariness may, in certain cases, be partly reduced by devices of motivation — more precisely relative motivation. On the one hand, he cites the notorious case of onomatopoeic words, such as Fr. *ouaoua* 'woof-woof'. This is the type of motivation foreshadowed in the *Cratylos* dialogue for the 'earliest words'. On the other hand, Saussure shows that formally complex words, such as Fr. *poirier* 'pear-tree', are relatively motivated by their morphological elements (cf. Fr. *poire* 'pear'). This is the type of motivation corresponding to the decomposition of 'later' words according to the *Cratylos* dialogue. Interestingly, Saussure already grasps the typological implications of the problem of lexical motivation (an issue we will come back to in Section 4).

In the present paper we re(de)fine the notion of lexical motivation on the basis of formal and cognitive parameters and present a pilot study analysing the 500 most frequent lexical words of French in order to get a provisional *motivational profile* of French high-frequency vocabulary. While we are aware of the role of motivation in grammar and at higher levels of language structure, we limit ourselves in the present paper to lexical motivation.

1.   The motivational square

One of the founding fathers of lexical typology, Stephen Ullmann stressed that problems of motivation form a parameter for linguistic typology.

> It seems clear [...] that no language is either completely transparent or completely opaque. All of them are likely to contain both conventional and 'motivated' terms in varying proportions [...] (Ullmann 1966: 221).

However, motivation is not only a quantitative, but also a qualitative typological variable. Inspired by Saussure (cf. section 0), but going beyond his preliminary reflections, Ullmann (1966: 221–222) distinguishes three types of motivation:

i. *Phonetic motivation* is surely a point that deserves attention in lexical typology. Certainly, all languages contain onomatopoeic words like Engl. *swish, sizzle, cuckoo*, as acknowledged by Saussure (for more detail, cf. Sharp and Warren 1994; Bredin 1996; Koch 2001a: 1157; Ungerer 2002: 372–376). But are there languages that are particularly prone to onomatopoeia and others that make less use of it? We will not explore this issue further, because we consider onomatopoeic words as being of secondary importance for the lexicon (cf. also Saussure 1916: 102). Our main concern will be the types of motivation considered under (ii) and (iii).

ii. *Morphological motivation* is due to word formation as in the following cases of suffixation (1a) and compounding (1b):

(1)   a.   Engl. *thinker*   (↔ *think*)
      b.   Engl. *armchair*   (↔ *arm*
                              ↔ *chair*)

This is the main type of 'relative motivation' Saussure identified in his *Cours de linguistique générale* (cf. Section 0).

iii. *Semantic motivation* is due to semantic relations between different senses of the same word (an innovation with respect to Saussure). In this context Ullmann adduces almost exclusively metaphors like the following:

(2)   a.   Engl. *bonnet* 'cover of a motor-car engine' (↔ *bonnet* 'head-dress')
      b.   Engl. *pivot* 'crucial point' (↔ *pivot* 'short shaft or pin on which something turns')

Although Ullmann admits that combinations of (ii) and (iii) may occur (1962: 92), he distinguishes (i), (ii) and (iii) in principle as three *types* of motivation. We shall argue, however, that the combination of (ii) and (iii) is not an accidental, but a systematic phenomenon of lexical motivation. We want to show that 'morphological' and 'semantic' motivation are not at all two distinct types, but two *dimensions* of non-onomatopoeic motivation that cross-classify. Let us consider some examples:

(3)   a.   Engl. *to open a bank account* 'to establish' (↔ *to open* 'to make open')
           Fr. *ouvrir un compte en banque* 'to establish' (↔ *ouvrir* 'to make open')

In these two languages, we would have, according to Ullmann, cases of semantic motivation, which, by the way, are perfectly parallel: Engl. *to open a bank account* is a transparent metaphor with respect to *to open a door* etc., as is Fr. *ouvrir*. The words *to open* and *ouvrir* display a (metaphorical) polysemy. In contrast to this, German makes use of prefixation to express the same concept:

(3)   b.   Germ. *ein Bankkonto eröffnen* 'to establish' (↔ *öffnen* 'to make open')

According to Ullmann's terminology, the German example would be a case of morphological motivation, since word formation, like prefixation, constitutes the essence of this kind of motivation. And yet, the semantic/cognitive relation between the two concepts expressed in examples (3a) and (3b) is exactly the same: the concept TO OPEN is mapped, in a metaphorical way, on the concept TO ESTABLISH (AN ACCOUNT). The only difference lies in the fact that English and French express this metaphor by means of polysemy of one and the same word, whereas German makes use of word formation. So, both (3a) and (3b) are examples of semantic/cognitive motivation.

Furthermore, in both (3a) and (3b) morphological/formal problems are involved. Clearly, prefixation, as in (3b), is a formal, more specifically a morphological device enabling speakers to create motivation. But consider (3a): *polysemy*, too, constitutes a *formal* device, enabling speakers to create motivation. It is surely a very special, extreme solution, but at the same time a very frequent one, and certainly a formal device, *viz.* total formal *identity*. So, the examples in (3a) contain not only cognitive, but also formal motivation. Similarly, (3b) is an example not only of formal, but also of cognitive motivation.

Let us give a further, more complex example. As noted above, Ullmann illustrates 'semantic' motivation nearly exclusively with metaphors, without denying, however, that other types of figurative expressions are also relevant for motivation. Indeed, one other important figurative device creating motivation is *metonymy* that is based on conceptual *contiguity*, the fundamental connection underlying all kinds of frames, scenarios, scripts etc. and including part-whole relations (cf. Fillmore 1975; 1985; Schank and Abelson 1977; Barsalou 1992; Taylor 1995: 87–92; Croft 1993; Ungerer and Schmid 1996: 205–217; Koch 1999a; Croft and Cruse 2004: 7–21). Table 1, which concerns the conceptual frame TREE, is centred on the contiguity relation TREE — FRUIT, and specifically on the contiguity relation PEAR-TREE — PEAR. This widespread, probably universal case of contiguity, which presents itself as a basis for linguistic motivation, can be expressed by very different formal devices in different languages of the world (cf. Koch 1999b).[1]

Summing up, we may say that formal and cognitive motivation are two dimensions of non-phonetic linguistic motivation, since there is no formal motivation without cognitive motivation and vice versa. So, we can replace Ullmann's binary classification by a "motivational square" (Figure 1) including both the formal and the cognitive component of lexical motivation (cf. also Marchand 1969: 2; Gauger 1971: 8; Rettig 1981: 21, 33–45, 209; Zöfgen 2002:

Table 1. Formal relations expressing the contiguity relation PEAR-TREE — PEAR (cf. Koch 1999b)

	C_1: PEAR-TREE	C_2: PEAR[a]
polysemy	Russ. *gruša*	*gruša*
	Czech I *hruška*	*hruška*
	Sard.I *pira*	*pira*
gender alternation	Ital. *pero*, m.	*pera*, f.
	Rum. *păr*, m.	*pară*, f.
	Lat. *pirum*, n.	*pirus*, f.
	Anc.Gr. *ápion*, n.	*ápios*, f.
suffixation	Fr. *poirier*	*poire*
	Span. *peral*	*pera*
	Port. *pereira*	*pera*
	Cat. *perer(a)*	*pera*
	Czech II *hrušeň*	*hruška*
	Mod.Gr. *axlaðjá, apiðjá*	*axláði, apíði*
(!)	Pol. *grusza*	*gruszka*
compounding	Engl. *pear-tree*	*pear*
	Germ. *Birnbaum*	*Birne*
	Swed. *päronträd*	*päron*
	Ndl. *pereboom*	*peer*
	Hung. *körtefa*	*körte*
	Jap. *nashinoki*	*nashi*
	Chin. *lí shù*	*lí*
	Breton. *gwez-pér*	*pér*
lexicalized syntagm	Turk. *armut ağacı*	*armut*
	Sard. II *arbore de pira*	*pira*
	Pers. *deraxt-e golābi*	*golābi*
	Arab. *šağara al-kummaθrai*	*kummaθrai*

[a] In the present context we focus on the element that is relevant for the contiguity relation involved. Of course every compound and lexicalised syntagm in Table 1 could be analysed according to its second element, as exemplified in Figure 4.

190).[2] A lexical item L_1 is motivated with respect to a lexical item L_2, if there is a cognitively relevant relation between the concept C_1 expressed by L_1 and the concept C_2 expressed by L_2 and if this cognitive relation is parallelled by a perceptible formal relation between the signifiers of L_1 and L_2.[3]
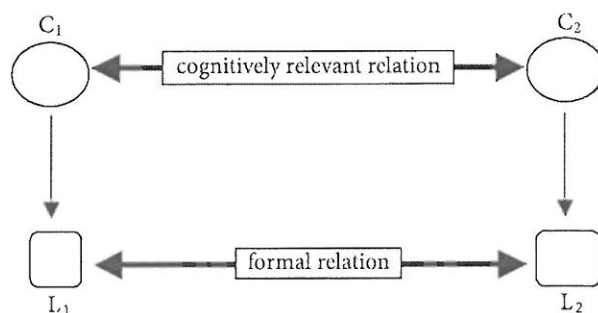
Figure 1. Motivational square (cf. Koch 2001a, 1156)

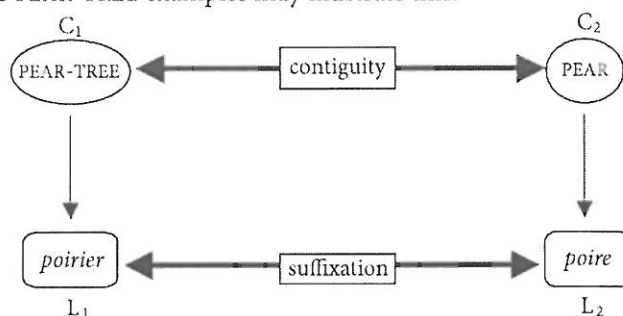One of our PEAR-TREE examples may illustrate this:

Figure 2. Motivational square, exemplified (Fr. *poirier* 'pear-tree'; cf. Table 1, suffixation)

The concept under examination, C_1, is PEAR-TREE. In French it is expressed by the lexical item $L_1 = poirier$, which is a motivated lexical item for two reasons:[4]

– There is a cognitively relevant relation, viz. contiguity between the concept $C_1 =$ PEAR-TREE expressed by L_1 *poirier* and the concept $C_2 =$ PEAR expressed by L_1 *poire*. The entities involved here are indeed concepts and not linguistic signifieds in Saussure's sense (1916: 97–100, 158–162). As shown by Benveniste (1966), the problem of arbitrariness and (relative) motivatedness of the linguistic sign does not concern the relation between the linguistic 'signifier' and 'signified' (which form, together, the linguistic sign and are therefore obligatorily bound by convention), but the relation between the linguistic sign and what Benveniste calls the 'thing'. Since motivation does not concern things as concrete referents in a given speech situation, but our general conceptual knowledge, we would rather speak of a 'concept' as a virtual entity (even in the case of onomatopoeia where stereotypical acoustic features of a sound phenomenon are imitated).

– There is a perceptible formal relation between the lexical item $L_1 = poirier$ and another lexical item $L_2 = poire$. We speak of 'lexical items' here, because it is not only the signifier in a narrow, phonological sense that is at stake here, but also its morphological structure. In the case of Fr. *poirier*, the formal relation is one of partial identity, because $L_2 = poire$ is part of $L_1 = poirier$. In other cases, as for instance word-class alternation, the formal relation involves the grammatical categories of L_1 and L_2 (cf. Table 3, ex. 70; Table 4, ex. (f)).

2. Particular cases: Polysemy, compounding, and absence of motivation

The very general Figure 1 can be adapted to particular cases of lexical motivation and can even demonstrate what absence of motivation means.

Obviously, in the case of polysemy the motivational square has to be reduced, because the lexical item L_1 is identical to L_2. Take, for instance, Russian *gruša*:
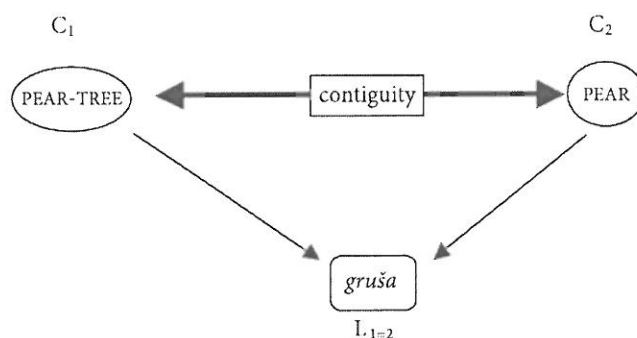
Figure 3. Reduced motivational square (Russ. *gruša* 'pear; pear-tree'; cf. Table 1, polysemy)

The cognitively relevant relation between C_1 and C_2 is (conceptual) contiguity, just like in Figure 2, but in this case, the formal relation is one of total identity of $L_1 = L_2 = gruša$.[5] This is a case of metonymic polysemy. Since in fact there is only one lexical item $L_{1=2}$, we can speak of *intrinsic* lexical *motivation* here, whereas the examples illustrated in Figures 2 and 4 are cases of *extrinsic* lexical *motivation* between different, though formally related, lexical items.[6]

In the case of compounds, as e.g. the lexical item L_1 = Hung. *körtefa*, the motivational square has to be made more complex, because L_2 and C_2 are doubled:

Figure 4. Complex instantiation of the motivational square (Hung. *körtefa* 'pear-tree'; cf. Table 1, compounding)
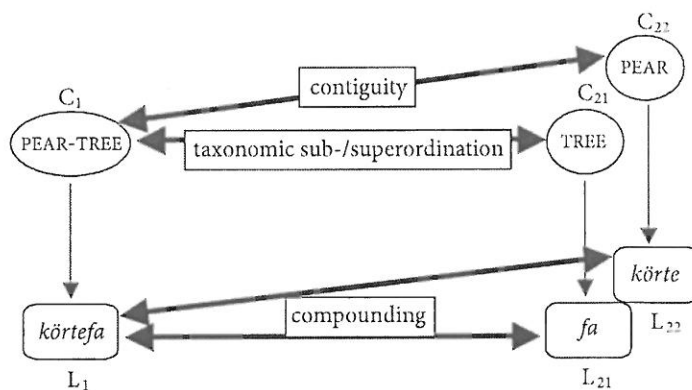
On the one hand, there is a cognitively relevant relation, namely taxonomic subordination, between the concept C_1 = PEAR-TREE and the concept C_{21} TREE as well as a perceptible formal relation between the lexical item L_1 = *körtefa* expressing C_1 and another lexical item L_{21} = *fa* expressing C_{21}. On the other hand, there is a cognitively relevant relation, namely contiguity, between the concept C_1 = PEAR-TREE and the concept C_{21} PEAR as well as a perceptible formal relation between the lexical item L_1 = *körtefa* expressing C_1 and another lexical item L_{22} = *körte* expressing C_{22}.

As the motivational square demonstrates, the concurrence of a cognitively relevant relation and a perceptible formal relation is the necessary condition for non-phonetic lexical motivation. If one of the two is lacking, there is no motivation. Figure 5 illustrates a case where there is a cognitively relevant relation without any formal relation:
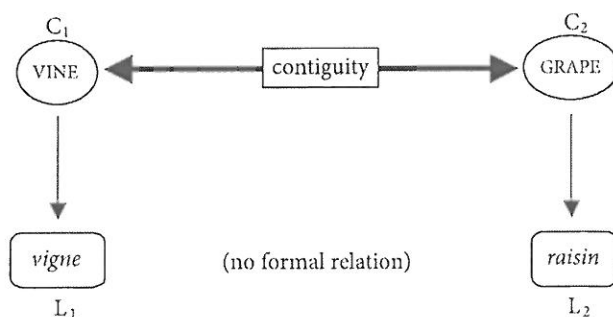
Figure 5. Absence of lexical motivation on formal grounds (Fr. *vigne*)

Between the concepts $C_1 =$ VINE and $C_2 =$ GRAPE, there is a clear cognitively relevant relation, viz. contiguity, but the two lexical items expressing them in French, $L_1 = vigne$ and $L_2 = raisin$ respectively, are totally unrelated on the formal level. So, we can speak of a lexically relevant relation (on cognitive grounds), but not of lexical motivation.

Figure 6, in contrast, illustrates a case where there is a perceptible formal relation without any cognitively relevant relation.
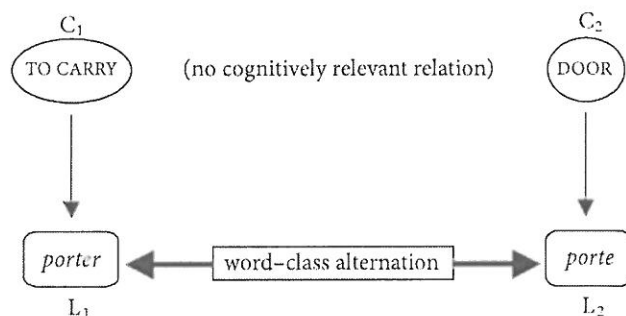
Figure 6. Absence of lexical motivation on cognitive grounds (Fr. *porter*)

Between the lexical items $L_1 =$ Fr. *porter* and $L_2 =$ Fr. *porte*, there seems to be a formal relation, because *porte* is homonymous to a part of $L_1 = porter$, but the concepts $C_1 =$ TO CARRY and $C_2 =$ DOOR, expressed by L_1 and L_2 respectively, are totally unrelated on the cognitive level. So, the partial formal identity is a mere accident. An extreme case of such "accidents" would be real lexical homonymy as between $L_1 =$ Fr. *son* 'bran' and $L_2 =$ Fr. *son* 'sound', with total formal identity and without any cognitively relevant relation.

Table 2. Non-phonetic lexical motivation — a two-dimensional grid (cf. Koch 2001a, 1160)

absence of mo-tivation	motivation							
	formal relations	cognitive relations						
		concep-tual identity	contigu-ity	meta-phorical similarity	cotaxo-nomic similarity	taxonomic superor-dination	taxonomic subordina-tion	con-ceptual contrast
	formal iden-tity/ polysemy	00	01	02	03	04	05	06
	tone alternation	10	11	12	13	14	15	16
	reduplication	20	21	22	23	24	25	26
	number alternation	30	31	32	33	34	35	36
	gender alternation	40	41	42	43	44	45	46
	voice alternation	50	51	52	53	54	55	56
	stem alternation	60	61	62	63	64	65	66
	word-class alternation	70	71	72	73	74	75	76
	suffixation	80	81	82	83	84	85	86
	prefixation	90	91	92	93	94	95	96
	compounding	100	101	102	103	104	105	106
	serial verb	110	111	112	113	114	115	116
	lexical. syntagm	120	121	122	123	124	125	126
	idiom	130	131	132	133	134	135	136
	...	...	...	...	...	...	...	...
	...	...	...	...	...	...	...	...

3. A two-dimensional grid for the description of lexical motivation

If we now want to describe the qualitative details of motivation (cf. Koch 2000a: 105–107; 2001a: 1156–1168), we have to specify, on the one hand, the different types of cognitively relevant relations C_1–C_2 and on the other hand, the different formal relations L_1–L_2. This will enable us to establish a two-dimensional grid like the one in Table 2, yielding a cross-classification of types of lexical motivation, where the horizontal axis corresponds to the cognitive relations (3.1) and the vertical axis to the formal relations (3.2). A supplementary aspect will be commented on in 3.3.

The numbers appearing in Table 2 (*viz.* 01, 02, etc., 10, 11, etc., 20, 21 etc.) are purely arbitrary and only serve as a means of identifying the different squares in the table.

3.1 Cognitive relations

The *cognitive* relations, corresponding to the horizontal axis in Table 2, ultimately go back to the three associative relations of 'contiguity', 'similarity', and 'contrast', that have been well established since Aristotle and have been corroborated by Husserl's phenomenology (cf. Holenstein 1972), by gestalt psychology (cf. Wertheimer 1922/23 and Raible 1981), and by free association tests (cf. Raible 1981), that have been introduced into linguistics by Roudet (1921: 686–692) and Jakobson (1956), and that have been applied with success to problems of lexical semantic change (cf., once more Roudet, and Ullmann 1962: 211–227, as well as the works cited in the following paragraph).

Now, lexical semantic change generates lexical motivation: forming a new word by derivation, compounding and the like ends up adding a new, motivated word to the vocabulary; creating a new sense for an already existing word ends up establishing motivation by polysemy; and so on (Koch 1991; 1994; 1999a; 2000b: 81–89; 2001b; 2001c: 14–25; in press; Blank 1997a; 1997b; 1998; 2000; 2001: 69–95, 104–108, 119–126; 2003a and b; Gévaudan 2003a and b). Consequently, the cognitive relations supporting semantic lexical change are likely to support lexical motivation, too.

By differentiating and combining contiguity, similarity, and contrast in various ways, we arrive at the universal, closed inventory of seven cognitive relations shown on the horizontal axis of Table 2 (for an additional, rather rare, type of contrast, at least in semantic change, cf. Blank 1997a, 220–225). They can be illustrated simply with some English examples of motivated words, neglecting for the moment the formal relations (cf. 3.2) and the problem of directionality (cf. 5.3.2):

(4) a. conceptual identity: Engl. *freedom* (↔ *free*)
     b. contiguity: Engl. *banker* (↔ *bank*)
     c. metaphorical similarity: Engl. *to grasp* 'to comprehend' (↔ 'to seize and hold')
     d. co-taxonomic similarity (i.e. taxonomic sisterhood): Engl. *bullock* (↔ *bull*)
     e. taxonomic superordination: Engl. *luck* 'chance' (↔ 'good fortune')
     f. taxonomic subordination: Engl. *luck* 'good fortune' (↔ 'chance')
     g. conceptual contrast: Engl. *impossible* (↔ *possible*)

Table 3. Formal (and cognitive) relations of motivation: examples from different languages

square No. in Table 2		formal relations	cognitive relations
absence of motivation	Engl. *queen*	—	(e.g. co-taxonomic similarity Engl. *king*)
01	Fr. *bois* 'woods' (↔ *bois* 'wood')	formal identity (= polysemy)	contiguity
		metonymic polysemy	
02	Engl. *to grasp* (cf. (4c))	formal identity (= polysemy)	metaphorical similarity
		metaphorical polysemy	
11	Chin. *mǎi* 'buy' (↔ *mài* 'sell')	tone alternation	contiguity
25	Réunion Creole *kozekoze* 'to chat' (↔ *koze* 'to speak')	reduplication	taxonomic subordination
31	Fr. *reins* 'loin' (↔ *rein* 'kidney')	number alternation	contiguity
45	Swahili *jitu* 'giant' (↔ *mtu* 'man')	gender alternation	taxonomic subordination
51	Anc.Gr. *misthoûsthai* 'to rent' (↔ *misthoûn* 'to let')	voice alternation	contiguity
65	Arab. *kassara* 'to smash' (↔ *kasara* 'to break')	stem alternation	taxonomic subordination
70	Ital. *rimborso* 'reimbursement' (↔ *rimborsare* 'to reimburse')	word-class alternation	conceptual identity
71	Chin. *shǎng* 'to proceed upwards' (↔ *shàng* 'up, above')	word-class alternation	contiguity
85	Ital. *libraccio* 'bad book' (↔ *libro* 'book')	suffixation	taxonomic subordination
96	Engl. *impossible* (cf. (4g))	prefixation	conceptual contrast
105 101	Germ. *Stadtrand* 'outskirts' (↔ [a] *Rand* 'edge'; ↔ [b] *Stadt* 'town')	compounding	[a] tax. subordination [b] contiguity
111 111	Yoruba *gbé ... wá* 'to bring' (↔ [a] *gbé* 'to carry'; ↔ [b] *wá* 'to come')	serial verb	[a] contiguity [b] contiguity
...	...	...	...

Freedom (4a) expresses fundamentally the same concept as *free* (though in a different word class: cf. the example Ital. *rimborso* in Table 3) (70). As for (4b), the concept C_1 = BANKER is linked to the concept C_2 = BANK by contiguity (AGENT–INSTITUTION within an ACTIVITY frame).[7] The concept C_1 = TO COMPREHEND expressed by *to grasp* stands in a relation of metaphorical similarity to the concept C_2 = TO SEIZE AND HOLD expressed by this same word (4c). The concept C_1 CASTRATED MALE BOVINE expressed by *bullock* and the concept C_2 = UNCASTRATED MALE BOVINE expressed by *bull* are parallel sub-categories within the taxonomic hierarchy subordinated to OX; so, we can speak of 'co-taxonomic similarity' (4d).[8] Still from the taxonomic point of view, the concept GOOD FORTUNE expressed by the word *luck* is subordinated to the concept FORTUNE expressed by this same word and, vice versa, the latter is superordinated to the former (4e, f).[9] The concepts C_1 = NOT POSSIBLE expressed by *impossible* and the concept C_2 = POSSIBLE expressed by *possible* are in contrast to each other within the same conceptual taxonomy.

3.2 Formal relations

The vertical axis of Table 2 corresponds to the *formal* relations involved in motivation. In contrast to the cognitive relations, the formal relations constitute an open inventory that has to account for the great variety of morpholexical devices fulfilling lexical motivation tasks in the world's languages (cf. Koch 2001a: 1159–1161). Table 3 contains some illustrations,[10] indicating, for each example, the corresponding square in Table 2 (first column), the lexical items involved and their meanings/senses (second column), the relevant formal relation (third column), and, additionally, the relevant cognitive relation(s) (fourth column). Note that the line labelled 'formal identity' in Table 2 corresponds from a synchronic-motivational perspective to polysemy according to Figure 3 (cf. the examples of type 01 and 02 in Table 3).

The types defined by the cross-classification of the horizontal and the vertical axis of Table 2 provide us with a descriptive tool for characterizing any kind of non-phonetic lexical motivation. Only the first example (Engl. *queen*) exemplifies the absence of motivation represented in the leftmost column of Table 2 (a rather common case in the lexicon!).

3.3 A further aspect: Formal transparency

In two diachronic lexicological projects at the University of Tübingen (cf. Blank et al. 2000; Gévaudan et al. 2003; contributions in Mihatsch and Steinberg 2004), we use a three-dimensional grid, including the two dimensions shown in Table 2, to describe lexical change. In these projects, a third dimension serves to capture the differences between autochthonous and borrowed lexical material. For a synchronic study dealing with motivation, the non-autochthonous character of lexical items is less interesting, or let us say: it is often subordinated to the problem of *formal transparency* (cf. Koch 2001a: 1162–1164). For example, Fr. *humain* is a learned word, i.e. a borrowing from Latin, from the diachronic point of view but what counts from the synchronic point of view is its reduced transparency with respect to Fr. *homme* (5a). On the other hand, we often encounter also reduced transparency without borrowing, as for instance in Fr. *jeu* with respect to Fr. *jouer* (5b).

(5) a. Fr. *humain* 'human' (↔ *homme* 'man, person')
     b. Fr. *jeu* 'game' (↔ *jouer* 'to play')

Thus, what is at issue here is the formal side of what Dressler (1985: 328–333) calls 'diagrammaticity'. Simplifying somewhat Dressler's eight-threshold scale, and introducing the phenomenon of polysemy as well, we can distinguish four degrees of formal transparency and one degree of non-transparency, i.e. non-motivation:

(6) Degrees of formal transparency:
     Ia = trivial cases of total transparency due to polysemy;
         e.g. Fr. *bois* 'woods' (↔ *bois* 'wood').
     Ib = total transparency: no allomorphic variation in the lexeme;
         e.g. Fr. <u>*service*</u> 'service' (↔ <u>*servir*</u> 'to serve').
     II = reduced transparency: allomorphic variation at that end of the
         lexeme where an affix is added;
         e.g. Fr. <u>*décision*</u> 'decision' (↔ <u>*décider*</u> 'to decide').
     III = minimal transparency: allomorphic variation in other parts of the
         lexeme;
         e.g. (5b) Fr. *jeu* 'game' (↔ *jouer* 'to play')
     IV = non-transparency = no motivation: strong suppletion etc.
         (equivalent to 'absence of motivation' in Table 2);
         e.g. Fr. *vite* 'quickly' (cf. *rapide* 'quick').

4.   Typological aspects

What is the typological relevance of the two-dimensional grid delineated in Section 3.? We discern three major scopes for its application:

(7)   (i)   motivational profiles of particular languages;
      (ii) identification of cross-linguistic motivational tendencies and idiosyncrasies with respect to the language facts (are there more or less formally transparent languages? Are there predominantly metaphorical languages? etc.)
     (iii) universal/language type-specific motivational preferences and gaps with respect to combinations in the two-dimensional grid in Table 2.

To begin with, the ultimate aim would be (7iii). We would like to know which combinations of cognitive relations, formal devices and — possibly — transparency qualities occur in different languages of the world and which ones are more or less "natural". Note that the two-dimensional grid is universally applicable:

–   The inventory of cognitive relations is universal and language-independent. As shown by phenomenological and gestalt psychological research as well as by historical semantics and as corroborated by empirical studies (cf. 3.1), there simply exist no other relations to connect conceptual contents.
–   The inventory of formal relations, an open list, can be enlarged in response to any typological characteristics. First of all, polysemy, as one of the formal relations of motivation (line 01–02–03 etc. in Table 2), is a universal. As for the other formal relations, we include any mechanism capable of producing a perceptible link interrelating the form of two lexical items in any language. As shown by the variety of examples in Table 3, the range of relevant phenomena includes devices going far beyond plain word formation, such as tone alternation, reduplication, gender alternation, voice alternation, stem alternation, serial verbs, etc.[11]
–   The scale of formal transparency is a cross-linguistically attested phenomenon, at least in inflectional languages, which are characterised among other things by phenomena of allomorphic variation in fusion (cf. Plungian 2001).
–   Nevertheless, Table 2 is only a heuristic grid. Table 3 gives just a sample of the possible combinations of cognitive and formal relations, and others do exist. However, not all the theoretically possible combinations represented in Table 2 are necessarily realized in any of the world's languages. We would like to know the universal or language type-specific preferences and

gaps and to explain them systematically, if possible. For example, we have so far found only examples of the types 70 and 71 in Table 2 (cf. Table 3 no. 70: Ital. *rimborso* 'reimbursement', and no. 71 Chin. *sháng* 'to proceed upwards'). Provided that further research confirms this fact, it could be easily explained: Metaphorical similarity and taxonomic relations between lexical units (similarity, super-/ subordination and contrast) seem to be incompatible with shifts of word classes. On the contrary, the "transpositional" nature of word-class alternation does not only — somewhat trivially — assure the identity of a given concept across word classes, but is also well adapted to words of different syntactic functions expressing contiguities within conceptual frames.

But before we come to such conclusions, we have to describe many different languages (7i). Even here, interesting insights are in store for us. Ferdinand de Saussure (1916: 183) already distinguished "langues lexicologiques", whose lexicon is maximally (though never totally) non-motivated, from "langues grammaticales", whose lexicon displays a considerable amount of motivatedness. Consequently, in traditional lexicology and typology, certain generalizations about lexical characteristics of particular languages are often mentioned. Stephen Ullmann, for instance, does not hesitate to claim, when looking at French:

> Nous avons vu la prédominance en français du mot simple et immotivé, dépouillé d'associations formelles (Ullmann 1969: 142).

The two-dimensional grid (Table 2) is a tool that allows us to test generalizations like these in a more precise manner, factorizing the problem of motivation according to different cognitive relations, different formal relations and degrees of formal transparency. Applied to central parts of the lexicon of a given language, it will provide us with a typological profile of lexical-motivational characteristics in this language (see (7i) above).

When we have described a certain number of widely differing languages (as different as possible!), we will be able to compare and contrast motivational profiles of various languages and to detect cross-linguistic tendencies and idiosyncrasies (see (7ii) above).

Table 4. Examples taken from the sample of the 500 most frequent French lexical words in Baudot 1992

	L_1	C_2	cognitive relation (cf. 3.1)	formal relation (cf. 3.2)	diagram-maticity (cf. 3.3)	L_2	C_2
(a)	*arriver* (02)	TO HAPPEN	metaphor. similarity	identity = polysemy	Ia	*arriver*	TO ARRIVE
(b)	*nombreux* (81)	NUMEROUS	contiguity	suffixation	Ib	*nombre*	NUMBER
(c)	*donner lieu* (132)+(131)	TO GIVE RISE	met. sim. + contiguity	idiom	Ib	*donner lieu*	TO GIVE RISE
(d)	*dire*	TO SAY	formal and cognitive opacity				
(e)	*effet*	BILL OF EXCHANGE	cognitive opacity	identity	(Ib)	*effet*	EFFECT
(f)	*intérêt* (70)	INTEREST	conceptual identity	word-class alternation	II	*intéresser*	TO INTER-EST
(g)	*vue* (80)	VIEW	conceptual identity	suffixation	III	*voir*	TO SEE
(h)	*se faire* (51)	TO GET DONE, TO HAPPEN	contiguity	voice alternation	Ib	*faire*	TO DO
(i)	*manières*	MANNERS	contiguity	number alternation	Ib	*manière*	MANNER, WAY
(j)	*découvrir* (96)&(91)	TO UNCOVER	conceptual contrast & contiguity	prefixation	Ib	*couvrir*	TO COVER
(k)	*femme* (04/05)	WIFE	tax. super-/ subord.	identity = polysemy	Ia	*femme*	WOMAN
(l)	*classe* (03)	(SCHOOL) CLASS	co-tax. similarity	identity = polysemy	Ia	*classe*	(SOCIAL) CLASS

5. Lexical motivation in French high-frequency vocabulary

5.1 Methodology of the study

Being Romance linguists, we begin with the languages we know best. We have begun to elaborate a motivational profile of central parts of the French vocabulary, and we present here our first results.

First of all, we have to define what we consider to be "central parts" of a vocabulary. The ideal starting point seems to be an onomasiological one. We would need a list of universal concepts, something like the Swadesh basic vocabulary list (1955; 1960), or, even better, something like Wierzbicka's collection of semantic primes (1996; cf. also the compromise inventory proposed by Goddard 2001: 9). However, apart from other problems, these collections are

not sufficiently large (the Swadesh list contains only 174 or 87 lexically relevant concepts, depending on different versions; Goddard's inventory comprises 19; Wierzbicka has 23 lexically relevant concepts). In order to cover a considerable part of the central vocabulary, we would need several hundred lexical units!

Thus, we opted for another solution and chose a frequency dictionary of French (Baudot 1992).[12] Why is a frequency dictionary acceptable for this task? When we start from the most frequent words in a language, we are sure to cover at least a large part of the most frequent concepts expressed in its speech community. In this way, the seemingly semasiological approach is justified, at least partly, on onomasiological grounds. Furthermore, if the number of lexical units is sufficiently large, the random character of the sample guarantees a sufficient statistical dispersion and thereby a certain degree of representativeness.

It is this random character of the sample that compensates for another one of its a priori deficiencies: A semasiological frequency list of lexical items necessarily disguises the phenomenon of polysemy, although this is a particularly important formal device of — intrinsic — motivation, that we cannot ignore under any circumstances (cf. the comment to Figure 3 and Table 2: formal identity). For French, we tested a provisional, but practicable solution based on the particular lexicographical options of the *Lexis* dictionary (Dubois 1975) that indirectly yields us important cases of polysemy present in the sample, without, however, producing a proliferation of senses.[13]

The sample embraces the 500 most frequent lexical words listed in Baudot 1992. Selecting only the lexical words in a narrow sense (since it is lexical motivation that we wanted to study) meant excluding grammatical words (*le/la*, *ce/cette*, *je*, *moi*, *qui?* etc., as well as prepositions and conjunctions like *de*, *avec*, *et*, *que*, *quand*, etc.), proper names and their derivates (*français*, *américain* etc.), numerals and quantifiers (*un*, *deux*, *premier*, *tout*, *plusieurs*, etc.). Since intrinsic motivation was included (representing numerous cases of polysemy within the 500 lexical words analysed), the sample contains about 936 cases that could be checked for motivational relations vs. opacity.

5.2 Results of the French study

The sample described in Section 5 yields the following results, exemplified in Table 4 and explained subsequently.[14]

5.2.1 *Motivation vs. opacity*
The examples in Table 4 are all cases of motivation, except (d) and (e).

Fr. *dire* (d) 'to say' is a typical case of what we might call *formal opacity*: since there is no formal relation that motivates this verb extrinsically, i.e. with respect to another French word, it is pointless to look for a cognitive relation (though there are, to be sure, numerous cognitive relations to other, formally non related words!). This type corresponds to 294 = 31% of our sample (Figure 7).

Fr. *effet* (e) 'bill of exchange' shows formal identity with *effet* 'effect', but at the synchronic level, there is no discernible cognitive relation. In other words: it is an example of homonymy. This type of *cognitive opacity* corresponds to 34 = 4% of our sample (Figure 7).

Thus, all in all, we get 608 = 65% of cases of *motivation* and 328 = 35% of cases of *opacity* (Figure 7). This is a first interesting result. Remember that we are dealing, for the moment, with top-frequency vocabulary that is supposed to be in general — and not only in French — less motivated by word-formation.

We will see later on (5.2.2 and 5.2.3) that it is the inclusion of polysemy into lexical motivation that produces totally different results.
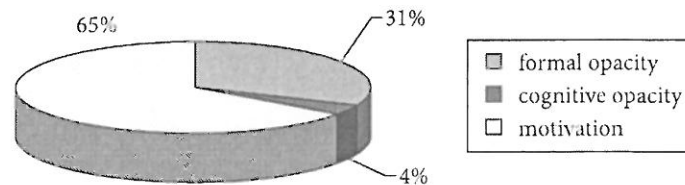

Figure 7. Percentages of motivation and opacity in the 500 most frequent lexical items of French (936 cases of motivation or opacity)

In the next subsections, we are going to sort out all the cases of motivation shown in Figure 7 and consider the internal composition of this group.

5.2.2   Formal transparency

Let us begin with formal transparency (cf. 3.3.; see the results in Figure 8 below). Cases of non-transparency (see IV. in (6) above) are systematically considered as non-motivated and, therefore, are no more included in Figure 8).

Example (g) in Table 4, Fr. *vue* 'view' in relation to *voir* 'to see', is a case of minimal transparency ((6), III). In our sample, there are very few cases of this type: 12 = 2% of all cases of motivation. Example (f) in Table 4, Fr. *intérêt* 'interest' in relation to *intéresser* 'to interest', is a case of reduced transparency ((6), II). Our sample contains relatively few cases of this type: 32 = 5% of all cases of motivation.
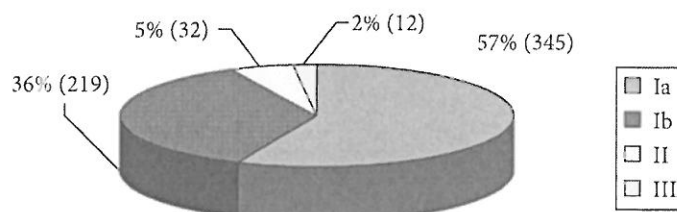
Figure 8. Percentages of degrees of formal transparency for 608 cases of motivation in the 500 most frequent lexical items of French (cf. Figure 7)

The remaining cases, all in all 564 = 93%, display total transparency. Within this huge percentage, we can distinguish two types.

On the one hand, we have all the cases of extrinsic motivation (cf. Section 2), i.e. different words that are related through word formation, grammatical devices, etc. and whose signifiers are partly or wholly identical. This holds for the following examples in Table 4: (b) Fr. *nombreux* 'numerous' in relation to *nombre* 'number', (c) Fr. *donner lieu* 'to give rise' in relation to *donner* 'to give' and *lieu* 'place', (h) Fr. *se faire* 'to get done, to happen' in relation to *faire* 'to do', (i) Fr. *manières* 'manners' in relation to *manière* 'manner, way', and (j) Fr. *découvrir* 'to uncover' in relation to *couvrir* 'to cover'. These cases of totally transparent extrinsic motivation ((6), Ib) make up 219 = 36% of our sample.

On the other hand, we have the cases of intrinsic motivation, i.e. polysemy (cf. Section 2). It is somewhat trivial to state that one and the same word form, expressing two related senses, has total formal transparency, but in some respect, it may be regarded as the highest degree of transparency ((6), Ia). Relevant examples are: (a) Fr. *arriver* 'to happen' and 'to arrive', (k) Fr. *femme* 'wife' and 'woman', and (l) Fr. *classe* 'form (at school)' and '(social) class'. These cases add up to 345 = 57%.

We are beginning to understand why motivation is present to so high a degree even in top-frequency vocabulary (Figure 7). This is due, in two thirds of all cases of motivation, to the presence of polysemy, which must not be neglected. Rather than being a French peculiarity, this may be a cross-linguistically recurrent pattern; but further research is necessary to confirm this (see also 7.1). We can only suppose for the moment that the percentage of polysemy will drop in the lower-frequency sections of vocabulary (cf. already Zipf 1949: 19–31), whereas the other types of motivation will increase, since subordinate level terms — as shown by prototype theory — are at the same time less frequent and morphologically more complex than basic level terms (Taylor 1995: 49; Ungerer and Schmid 1996: 88–92; Croft and Cruse 2004: 86).

5.2.3   Formal relations

Turning now to the formal dimension corresponding to the vertical axis of the two-dimensional framework in Table 2, we see that the results (see Figure 9 below) confirm what we already detected in the realm of formal transparency: in the present top-frequency sample, the overwhelming majority of cases of motivation, namely three fifths (361 = 59.6%), are due to polysemy, i.e. formal identity. We already cited the relevant examples (a) Fr. *arriver*, (k) Fr. *femme*, and (l) Fr. *classe* in Table 4.
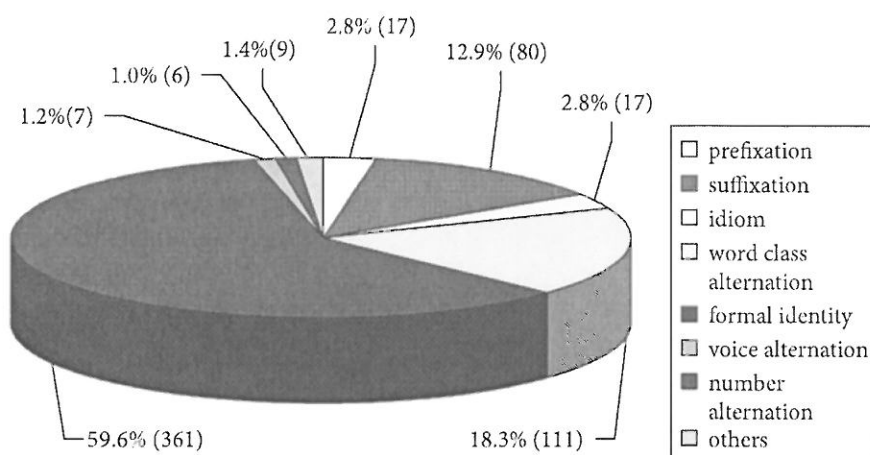

Figure 9. Percentages of formal relations in the 500 most frequent lexical items of French (608 cases of motivation)

The second most important formal device is word-class alternation (111 = 18.3%), exemplified by (f) Fr. *interêt* ↔ *intéresser* in Table 4. Another relatively frequent device is suffixation (80 = 12.9%), exemplified by (b) Fr. *nombreux* and (g) Fr. *vue*. Less frequent devices are: idioms (17 = 2.8%; cf. (c) Fr. *donner lieu*), prefixation (17 = 2.8%; cf. (j) Fr. *découvrir*), number alternation (6 = 1.0%; cf. (i) Fr. *manières*), voice alternation (7 = 1.2%; cf. (h) Fr. *se faire*), etc.

With respect to the Ullmann quote in Section 4., we have to observe that it is only the exclusion of polysemy that seems to make French a rather "unmotivated" language. As soon as we include polysemy, as our motivational square and even Ullmann's concept of 'semantic' motivation suggest that we do (cf. Section 1), the picture changes considerably (but cf. also 7.1).

5.2.4   Cognitive relations

Our last and most important issue, corresponding to the horizontal axis of the two-dimensional framework in Table 2, concerns the variety of cognitive

relations that underlie the different formal relations including polysemy (see Figure 10 below).

As already mentioned in Section 1, Ullmann illustrated what he called 'semantic motivation' mainly by metaphorical examples, and we think everybody would predict that it is metaphor that predominates in lexical motivation (cf. Lakoff and Johnson 1980). Indeed, metaphorical similarity, exemplified in Table 4 by (a) Fr. *arriver* 'to happen' and 'to arrive', *is* important, amounting to almost one fifth (113 = 18.6%) in our sample.

Nevertheless, the majority of all cases of motivation is based on contiguity (350 = 57.4%): e.g. Table 4, (b) Fr. *nombreux* 'numerous' in relation to *nombre* 'number'; (c) Fr. *donner lieu* 'to give rise' in relation to *lieu* 'place' (contiguity between the PLACE and the EFFECT realized in this place); (h) Fr. *se faire* 'to get done, to happen' in relation to *faire* 'to do', (i) Fr. *manières* 'manners' in relation to *manière* 'manner, way' (contiguity between MODE and CONDUCT, as expressible in 'mode of conduct').[15]
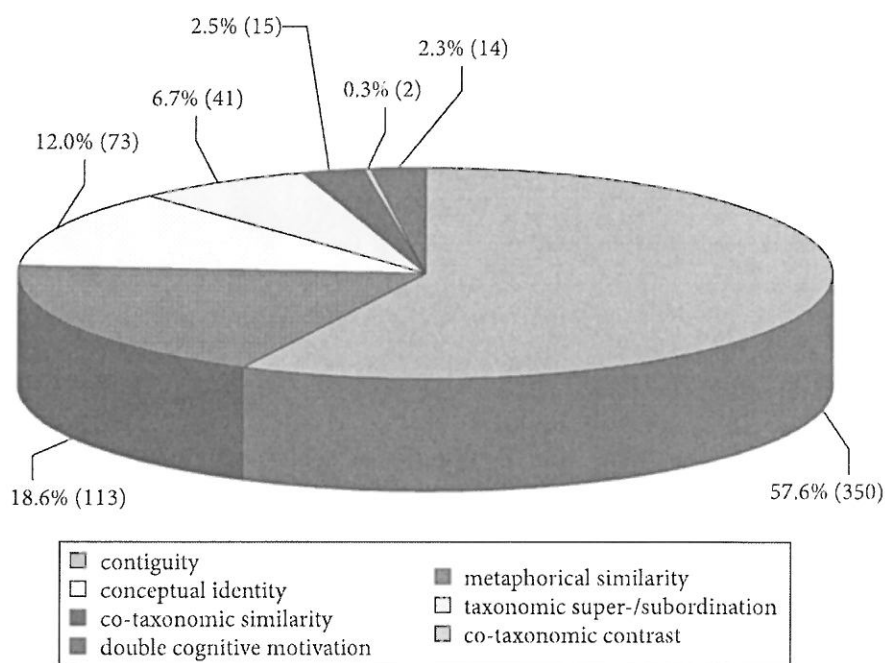

Figure 10. Percentages of cognitive relations in the 500 most frequent lexical items of French (608 cases of motivation)

Another relatively important relation is conceptual identity, typical of word-class alternation (73 = 12.0%; e.g. Table 4, (f) Fr. *intérêt* 'interest' in relation

to *intéresser* 'to interest'), but realized also through suffixation: e.g. (g) Fr. *vue* 'view' in relation to *voir* 'to see'.

Taxonomic relations, too, have some impact on lexical motivation, especially taxonomic super- and subordination realized through polysemy as in (k) Fr. *femme* 'wife' and 'woman'. The overwhelming majority of occurrences of super- and subordination are cases of synchronic polysemy such as this one, where the direction of intrinsic motivation is undecidable (cf. n. 9), and so we cannot reasonably separate these two relations. All in all, they correspond to 41 = 6,7% of our sample. Furthermore, we have some cases (15 = 2.3%) of co-taxonomic similarity, exemplified in Table 4 by (l) Fr. *classe* 'class (at school)' and '(social) class' (two different types of hierarchical divisions).

The cognitive profile of our sample constitutes the core of this approach. Thanks to the two-dimensional framework that factorizes lexical motivation, we are able to sort out the cognitive "finger-prints" of motivation in a given sample and in a given language.

The high percentage of contiguity surely is not fortuitous. Several studies in diachronic lexicology revealed the wide range and the frequency of metonymic change and in general of contiguity effects (cf. Blank 1997a: 230–231; Waltereit 1998; Koch 1999a; 1999b; 2001b, 2004). As already observed in Section 3.1, lexical motivation is generated by lexical change. Thus, they are likely to work according to the same principles and proportions. In fact, for contiguity there seem to be comparable proportions in diachrony and in synchrony. (see also 5.3.5).

5.3   Discussion of the French study

The results set forth in Section 5.2 represent a first step towards a systematic evaluation of lexical motivation in French high-frequency vocabulary. From the perspective of a typological research project as well as from the perspective of an extension to less-than-top-frequency vocabulary, they raise several questions.

5.3.1   *The problem of polysemy*
In Section 5.1, we opted for a provisional solution that permitted us to integrate polysemy in the a priori semasiological frequency list selected from Baudot 1992. Thanks to the special nature of the *Lexis* dictionary, this procedure prevented us from triggering a proliferation of senses (cf. n. 13). Even on this basis, however, we found ourselves compelled to exclude from the count certain senses of several words, because they were highly specialized and, to

be sure, not very frequent (belonging to languages for special purposes or the like), as for instance (8):

(8)   Fr. *point* 'stitch' ($\leftrightarrow$? *point* 'point')

But from which threshold on are we allowed to disregard senses appearing in our material, as it has been defined by our methodology? Whereas in the case of extrinsic motivation every single word is selected or not depending on its frequency, the problem is that we have no real control over the frequency of different *senses* of a word. It may be that the results concerning polysemy (Figures 8 and 9) exaggerate the weight of polysemy, because frequent *words* ordinarily are highly polysemous (cf. 5.2.3) — surely with additional differences between word classes (Fellbaum 1990). Thus, the selection of senses was not restrictive enough. We should base our study of high-frequency vocabulary rather on the frequency or salience of *lexical units* (embracing one form and one sense; cf. Cruse 1986: 49, 80; also Lipka 2002: 148–150) than on the frequency of words. In this respect, another method has to be elaborated, all the more as the procedure based on the dictionary *Lexis* (Dubois 1975) is not applicable to languages other than French because of the lexicographic peculiarities of this dictionary. A possible alternative consists in systematically questioning informants for the main senses of the sample words (in order to get sufficiently autonomous senses: cf. n. 13) before investigating the motivational relations. In this way, we are switching from linguist-based observations to native-speaker-based intuitions. In contrast to the procedure based on *Lexis*, this approach, which we are developing at the moment,[16] is easily replicable and thus transferable to any other language.

5.3.2   *The directionality of motivational relations*
In a case like Table 4, (b) Fr. *nombreux*, the motivational relation seems to be clearly directed: it is *nombreux*, as a derivative, that is motivated by *nombre*, and not vice versa. But the direction of motivation is not always so obvious (cf. Rettig 1981, 167–172). At any rate, there are two types of formal devices that raise problems.

First, in the case of polysemy a synchronic analysis cannot in general determine the direction of motivation (the diachronic direction of change, although we ordinarily know it, is totally irrelevant!), as is shown e.g. by (4e) and (4f) Engl. *luck* and Table 4, (k) *femme*. To standardize the methodology, we systematically treated intrinsic motivation (within polysemy) as non-directed, but this solution is not always satisfying, especially in cases like Table 4, (a) *arriver*, where we would expect TO HAPPEN to be metaphorically motivated by

TO ARRIVE and not vice versa. Even for cases like (4e/f) *luck* and Tabel 4, (k) *femme*, the non-directionality cannot be taken for granted.

Second, in the case of word-class alternation, the directionality of motivation often cannot be determined: is Fr. *intérêt* motivated by *intéresser* or the reverse? To standardize the procedure, we systematically treated the extrinsic motivation by word-class alternation as non-directed, but once more this solution is not always satisfying, as illustrated by the following examples:

(9)   Fr. *regarder* 'to look, to watch' — *regard* 'look'

(10) Fr. *peiner* 'to try hard, to struggle with s.th.' — *peine* 'pains'

Intuitively, we would consider (9) *regard* to be motivated by *regarder*, (10) *peiner* to be motivated by *peine*, but according to which criteria? The decision is not irrelevant, because with non-directionality (9) *regarder* would be motivated as well, with directionality either *regarder* or *regard* would not, and so on.

The only way to resolve the question if there is any directionality in polysemy and in word-class alternation is the integration of other types of data, by preference judgments of native speakers, which implies the elaboration of a special method of inquiry (cf. n. 16).

5.3.3   *Motivational relations: empirical bases*

Once we are ready to integrate speaker judgments (5.3.2), we might be able to revise the methodology even further. In the present research, it is the linguist who, in the end, makes the decisions on whether there is or is not a motivational relation between two lexical units (and which is the formal and which is the cognitive relation at issue). It would be tempting and empirically more sound to get this information from speakers, but we would still have to elaborate a rather sophisticated method of inquiry to elicit them.

5.3.4   *Cognitive transparency*

Another important aspect of the two-dimensional framework (Table 2) is formal transparency (3.3). On closer inspection — and as shown by the examples (5) and (6) — formal transparency is a kind of sub-dimension of the formal relations represented as vertical dimension. Yet, the graduality of transparency is not only a formal, but also a *cognitive* problem (cf. also Rettig 1981: 157–166). Words such as (b) *nombreux*, (f) *intérêt*, (g) *vue*, and (j) *découvrir* in Table 4 seem to be unproblematic from the point of view of cognitive transparency. But what about (i) *manières* in relation to *manière* or (h) *se faire* in relation to *faire*? And what about (c) *donner lieu* in relation to *donner* and above all to

lieu? In addition to the transparency aspect of the formal relation, we probably have to introduce a supplementary transparency sub-dimension of the cognitive relations. As the examples show, formal and cognitive transparency are not necessarily intertwined with each other, i.e. they seem to be independent parameters.

The degree of formal transparency can be decided on by the linguist according to purely formal, mechanical criteria (see (6) above). Cognitive transparency is different, all the more so as the linguist always runs the risk of applying etymological criteria that are out of place in a synchronic investigation. On the cognitive level too, it would be worthwhile, then, to develop a method to "measure" the degrees of transparency on the basis of the speakers' judgment.

6. Conclusion

This paper has re(de)fined non-phonetic lexical motivation, postulating that it *always* has a cognitive and a formal dimension at the same time. Starting from these premises, we proposed a two-dimensional grid comprising a closed universal inventory of cognitive relations and an open inventory of formal relations (specified by degrees of formal transparency). Thanks to this framework it will be possible to describe the motivational properties of lexical units in any language.

To give a sense of what such a motivational profile could look like, we presented a pilot study of lexical motivation in French high-frequency vocabulary. The most outstanding result concerns the relative proportions of cognitive relations: contrary to common expectations, it was not metaphoric but contiguity-based motivation that prevailed, whose percentage is overwhelming across all the different formal relations (followed, though at a great distance, by metaphorical similarity).

Independently of the problems and of the possible methodological and theoretical improvements discussed in 5.3., it will be most interesting to observe whether the percentages presented in Figures 7–10 will remain constant or will change, when we extend the sample — extending it, first, within French, to less-than-top frequency sections of vocabulary, and, second, to other languages. Surely, there will be differences as to formal relations (5.2.3) and formal transparency (5.2.2), perhaps also as to motivation and opacity in general (5.2.1). But what about the relative percentages of cognitive relations and especially the predominance of contiguity (5.2.4)? The comparison of cognitive profiles of lexical motivation in different languages will be a fascinating

issue: Do different languages have different cognitive "finger-prints" or do the (provisional) percentages observable in Figure 10 recur across languages? In order to be able to answer these and other questions, we are currently working out (cf. n. 16) a universally applicable method of investigation into lexical motivation based on native speaker judgments.

Notes

* We express our gratitude to Sam Featherston and Steven Dworkin for the stylistic revision of this paper.

1. While in the majority of cases, suffixation derives the word for PEAR-TREE from the word for PEAR, interestingly Polish has chosen the reverse direction. — The two types 'compounding' and 'lexicalized syntagm' may be realized either with the order of modifier-head (first subgroup in both groups) or with the order head-modifier (first subgroup in both groups). In Tables 1 and 2 (and throughout this paper) these two order possibilities are treated as just two variants of compounds (and lexicalised syntagms), whereas prefixations and suffixations, which represent two order possibilities as well, are systematically separated. This is a merely conventional decision due to current practice in word formation research. It may be easily revised on cross-linguistic grounds.

2. Note that the view on motivation symbolized in Figure 1 is restricted to lexical issues concerning form-form relations correlated with content-content relations (for a broader view of problems of "iconicity", "motivation", and "isomorphism" cf. Radden and Panther 2004 and Ungerer 2002).

3. Strictly speaking, motivation corresponds to something we should rather call a potential of 'motivatability' than a static condition of 'motivatedness' (Gauger 1971: 12, 152; Rettig 1981: 75–78). Nevertheless, the metalinguistic capability of motivation intervenes constantly even in ordinary language use, as is shown by phenomena such as understanding of (complex) lexical items hitherto unknown to the hearer, misunderstandings (deviant motivation), argumentation on lexical grounds, and puns (cf. Rettig 1981: 173–208). Last but not least, we should mention here the not very frequent, but highly significant phenomenon of popular etymology (cf. Blank 1997a: 303–317; 2001: 91–92). It just helps to remedy a situation of non-motivatedness like the one represented in Figure 6.

4. There is a certain analogy between our motivational square and the square Langacker (1987: 84) uses to represent the way in which an "integration of components at the semantic pole corresponds to, and is symbolized by, the integration of components at the phonological pole." Although centred on complex lexical expressions, our cognitively relevant relation specifies what Langacker, for grammatical constructions, calls 'integration of semantic components' and our formal relation roughly corresponds to the 'integration of phonological components'. The reason why we consider not only phonological but also morphological relations on the formal level will be explained below.

5. Schwarze (2001: 76) and Radden and Panther (2004: 20–21), likewise underline the motivational basis of polysemy.

6. This terminological distinction must not be confused with another one suggested by Guiraud (1955: 17–18.; cf. also Gusmani 1984: 15, and Zöfgen 2002: 189): 'internal' or 'intra-linguistic motivation' between signs of one and the same language (formal motivation of type (ii) and, possibly, cognitive motivation of type (iii)) vs. 'external' or 'natural motivation' by the extralinguistic reality (roughly, type (i)).

7. For the utilization of activity frames in word formation, cf. Schwarze 1995: 500–506; Blank 1998: 13–14.

8. Note that metaphorical similarity and co-taxonomic similarity can be clearly distinguished. Metaphorical similarity is based on domain-mapping (Croft 1993, Croft and Cruse 2004) i.e. conceptual leaps from one frame to another or from one taxonomy to another, whereas co-taxonomic similarity is bound to one and the same taxonomy (cf. Blank 1997a: 207–216).

9. From the synchronic point of view, neither of the two senses of *luck* is prior to the other. Cf. also 5.3.2.

10. We do not exemplify all formal relations present in Table 2, because it is an open list anyway (cf. Section 4).

11. Many formal categories interact with grammar, such as e.g. compounds, lexicalised syntagms and idioms. However, grammar is not in the scope of our interest, because we need to make a distinction between three different aspects: (i) the motivation of grammatical structure, which is not a lexicological problem, as in the case of iconicity in word order; (ii) the diachronic "motivation" of innovations with (naturally grammatical) combinations of lexical units in the formation of new lexical units, for example with compounds, lexicalised syntagms, idioms, etc.; (iii) the motivation of lexical units from a synchronic perspective. All of these three issues are of great importance and fully legitimate, but our synchronically and lexicologically oriented paper is necessarily limited to the third issue. This does not exclude — from a different perspective — a systematic relationship between (ii) and (iii). Obviously, in any given language, linguists would have to check for each of the formal processes to what extent it has already arrived, beyond the level of free grammatical composition, at stage (iii), and whether the corresponding process would have to be inserted in the open inventory of the vertical dimension of Table 2. At the same time, these considerations hold for compounds, lexicalised syntagms, idioms as well as for incorporations in polysynthetic languages (cf. Mithun 1986, Gerdts 1998).

12. Compared to other word frequency dictionaries — like C.N.R.S. NANCY 1971, Gougenheim et al. 1964, and Juilland et al. 1970 — this dictionary seemed to be the best available source for our purpose: its corpus basis (texts from 15 different genres written between 1906 and 1967 mainly in France and Canada) is more recent than the other dictionaries' and Baudot's way of dealing with phenomena like homographs, compounds and different word forms is much more transparent than the others'.

13. As for the criteria that can serve to determine the degree of 'autonomy' of different senses of a word, cf. Croft and Cruse 2004, 74–106. These very sophisticated criteria are highly interesting for detailed studies of single words, but they are too "costly" for a large-scale study like the present one. Thus, we had to opt for a methodological "shortcut" (but cf. 5.3.1) that takes advantage of the particular lexicographic approach characterizing the *Lexis* dictionary. This approach consists in increasing the homonymic split of entries. These multiple "homonymies" hide, in fact, many cases of what we would call polysemy, because there is a cognitively relevant relation between the concepts expressed according to Figure 3 (obviously in some other cases, there is "real" homonymy). Since *Lexis* could not infinitely maximize the homonymic split (and, thus, accounts for fine-grained polysemy *within* its entries), it offers the opportunity to limit the proliferation of senses, if we read separated "homonymic" entries as the main senses within a polysemy, inasmuch as they are connected by cognitive relations.

14 . In the second column, the number in brackets indicates the square number according to Table 2, except in cases of absence of motivation.

15. Example (j) Fr. *découvrir* belongs to a very special type of verbs that imply two motivating cognitive relations at the same time (in this case conceptual contrast & contiguity). TO UNCOVER and to TO COVER (Fr. *couvrir*) are connected not only by contiguity (inasmuch as the action of UNCOVERING presupposes a preceding action of COVERING or at least its result), but also by conceptual contrast, since they are opposites of each other. Thus, *découvrir* is one of the 14 cases (2.3%) in our sample that combine two cognitive relations (see Figure 10).

16. In the research project LexiTypeSyn (project B6 within the Collaborative Research Centre 441) at the University of Tübingen (http://www.sfb441.uni-tuebingen.de/b6/) we are currently investigating lexical motivation in selected portions of French, Italian and German vocabulary.

References

Barsalou, Laurence W. 1992. Frames, concepts, and conceptual fields. *Frames, fields, and contrasts. New essays on semantic and lexical organization*, Adrienne Lehrer and Eva. F. Feder Kittay (eds.), 21–74. Hillsdale NJ: Lawrence Erlbaum.

Baudot, Jean. 1992. *Fréquences d'utilisation des mots en français écrit contemporain*. Montréal: Presses de l'Université de Montréal.

Benveniste, Emile. 1966. Nature du signe linguistique. *Problèmes de linguistique générale* (1), 49–55. Paris: Gallimard.

Blank, Andreas. 1997a. *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*. Tübingen: Niemeyer.

Blank, Andreas. 1997b. Outlines of a cognitive approach to word-formation. *Proceedings of the 16th International Congress of Linguists*, CD-Rom Paper No. 0239.

Blank, Andreas. 1998. Kognitive italienische Wortbildungslehre. *Italienische Studien* 19: 5–27.

Blank, Andreas. 2000. Pour une approche cognitive du changement sémantique lexical: Aspect sémasiologique. Société de Linguistique de Paris 2000: 59–73.

Blank, Andreas. 2001. *Einführung in die lexikalische Semantik für Romanisten.* Tübingen: Niemeyer.

Blank, Andreas. 2003a. Polysemy in the lexicon and in discourse. *Polysemy. Flexible patterns of meaning in mind and language*, Brigitte Nerlich et al. (eds.), 267–293. Berlin: Mouton de Gruyter.

Blank, Andreas. 2003b. Words and concepts in time: Towards diachronic cognitive onomasiology. *Words in time. Diachronic semantics*, Regine Eckardt; Klaus von Heusinger, and Christoph Schwarze (eds.), 37–65. Berlin: Mouton de Gruyter.

Blank, Andreas; Peter Koch and Paul Gévaudan. (2000). Onomasiologie, sémasiologie et l'étymologie des langues romanes: Esquisse d'un projet. *Actes du XXIIᵉ Congrès International de Linguistique et de Philologie Romanes, Bruxelles, 23–29 juillet 1998. IV : Des mots aux dictionnaires*, Annick Englebert; Michel Pierrard; Laurence Rosier and Dan Van Raemdonck (eds.), 103–114, Tübingen: Niemeyer.

Bredin, Hugh. 1996. Onomatopoeia as a figure and a linguistic principle. *New Literary History* 27: 555–569.

Coseriu, Eugenio. 2004. *Der Physei-Thesei-Streit. Sechs Beiträge zur Geschichte der Sprachphilosophie.* Tübingen: Narr.

C.N.R.S. NANCY. 1971 (ed.). *Dictionnaire des fréquences. Vocabulaire littéraire des XIXe et XXe siècles.* Paris: Didier.

Croft, William. 1993. The role of domains in the interpretation of metaphors and metonymies. *Cognitive Linguistics* 44: 335–370.

Croft, William and D. Alan Cruse. 2004. *Cognitive linguistics.* Cambridge: CUP.

Cruse, D. Alan 1986. *Lexical semantics.* Cambridge: CUP.

Dressler, Wolfgang U. 1985. On the predictiveness of natural morphology. *Journal of Linguistics* 21: 321–337.

Dubois, Jean 1975. *Lexis. Dictionnaire de la langue française.* Paris: Larousse.

Fellbaum, Christiane 1990. English verbs as a semantic net. *International Journal of Lexicography* 3: 278–301.

Fillmore, Charles J. 1975. An alternative to checklist theories of meaning. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 1: 123–131.

Fillmore, Charles J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* 6: 222–254.

Gauger, Hans-Martin 1971. *Durchsichtige Wörter. Zur Theorie der Wortbildung.* Heidelberg: Winter.

Gerdts, Donna B. 1998. Incorporation. *The handbook of morphology*, Andrew Spencer and Arnold M. Zwicky (eds.), 84–100. Oxford: Blackwell.

Gévaudan, Paul. 2003a. Klassifikation lexikalischer Entwicklungen. Semantische, morphologische und stratische Filiation. PhD dissertation, University of Tübingen.

Gévaudan, Paul. 2003b. Lexikalische Filiation. *Kognitive romanische Onomasiologie und Semasiologie*, Andreas Blank and Peter Koch (eds.), 189–211. Tübingen: Niemeyer.

Gévaudan, Paul; Peter Koch and Antonia Neu 2003. Hundert Jahre nach Zauner. Die romanischen Namen der Körperteile im DECOLAR. *Romanische Forschungen* 115: 1–27.

Goddard, Cliff. 2001. Lexico-semantic universals: A critical overview. *Linguistic Typology* 5: 1–65.

Gougenheim, Georges et al. 1964. *L'élaboration du français fondamental (1er degré)*. Paris: Didier.

Guiraud, Pierre. 1955. *La sémantique*. Paris: Presses Universitaires de France.

Gusmani, Roberto. 1984. A proposito della motivazione linguistica. *Incontri linguistici* 9: 11–23.

Holenstein, Elmar. 1972. *Phänomenologie der Assoziation. Zur Struktur und Funktion eines Grundprinzips der passiven Genesis bei E. Husserl*. Den Haag: Nijhoff.

Jakobson, Roman. 1956. Two aspects of language and two types of aphasic disturbances. *Fundamentals of language*, Roman Jakobson and Morris Halle ²1971, 67–96. Den Haag: Mouton.

Juilland, Alphonse; Dorothy Brodin and Cathrin Davidovitch 1970. *Frequency dictionary of French words*. Den Haag: Mouton.

Koch, Peter. 1991. Semantische Valenz, Polysemie und Bedeutungswandel bei romanischen Verben. *Connexiones Romanicae. Dependenz und Valenz in romanischen Sprachen*, Peter Koch and Thomas Krefeld (eds.), 279–306, Tübingen: Niemeyer.

Koch, Peter. 1994. Gedanken zur Metapher — und zu ihrer Alltäglichkeit. *Sprachlicher Alltag. Linguistik — Rhetorik — Literaturwissenschaft. Festschrift für Wolf-Dieter Stempel 7. Juli 1994*, Annette Sabban and Christian Schmitt (eds.), 201–225, Tübingen: Niemeyer.

Koch, Peter. 1999a. Frame and contiguity. On the cognitive bases of metonymy and certain types of word formation. *Metonymy in language and thought*, Klaus-Uwe Panther and Günter Radden (eds.), 139–167, Amsterdam: John Benjamins.

Koch, Peter. 1999b. Tree and fruit. A cognitive-onomasiological approach. *Studi Italiani di Linguistica Teorica e Applicata* 28: 331–347.

Koch, Peter. 2000a. Indirizzi cognitivi per una tipologia lessicale dell'italiano. *Italienische Studien* 21: 99–117.

Koch, Peter. 2000b. Pour une approche cognitive du changement sémantique lexical: Aspect onomasiologique. Société de Linguistique de Paris 2000: 75–95.

Koch, Peter. 2001a. Lexical typology from a cognitive and linguistic point of view. *Language typology and language universals/Sprachtypologie und sprachliche Universalien/ La typologie des langues et les universaux linguistiques. An international handbook/ Ein internationales Handbuch/Manuel international*, Martin Haspelmath; Ekkehard König; Wulf Oesterreicher and Wolfgang Raible (eds.), II, 1142–1178. Berlin: Walter de Gruyter.

Koch, Peter. 2001b. Metonymy: Unity in diversity. *Journal of Historical Pragmatics* 2: 201–244.

Koch, Peter. 2001c. Bedeutungswandel und Bezeichnungswandel. Von der kognitiven Semasiologie zur kognitiven Onomasiologie. *Zeitschrift für Literaturwissenschaft und Linguistik* 121: 7–36.

Koch, Peter. 2004. Metonymy between pragmatics, reference and diachrony. *Metaphorik.de* 7: 6–54.

Koch, Peter. 2005. Taxinomie et relations associatives. *Sens et réferences. Mélanges pour Georges Kleiber/Sinn und Referenz. Festgabe für Georges Kleiber*, Adolfo Murguía (ed.), 159–191. Tübingen: Narr.

Lakoff, George and Mark Johnson. 1980. *Metaphors we live by*. Chicago IL: Chicago University Press.

Langacker, Ronald. 1987 and 1990. *Foundations of cognitive grammar*, 2 Vols. Stanford CA: Stanford University Press.

Lipka, Leonhard. 2002. *English lexicology. Lexical structure, word semantics, and word-formation*. Tübingen: Narr.

Marchand, Hans. ²1969. *The categories and types of present-day English word-formation. A synchronic-diachronic approach*. München: Beck.

Mihatsch, Wiltrud and Reinhild Steinberg (eds.). 2004. *Lexical data and universals of semantic change*. Tübingen: Stauffenburg.

Mithun, Marianne. 1986. On the nature of noun incorporation. *Language* 62: 32–38.

Plato. 1996. Cratylos. *Plato*, engl. transl. by Harold N. Fowler, IV, 3–191. Cambridge MA: Harvard University Press.

Plungian, Vladimir. 2001. Agglutination and flection. *Language typology and language universals/Sprachtypologie und sprachliche Universalien/La typologie des langues et les universaux linguistiques. An international handbook/Ein internationales Handbuch/Manuel international*, Martin Haspelmath; Ekkehard König; Wulf Oesterreicher and Wolfgang Raible (eds.), I, 669–678. Berlin: Walter de Gruyter.

Radden, Günter und Klaus-Uwe Panther. 2004. Introduction: Reflections on motivation. *Studies in linguistic motivation*, Klaus-Uwe Panther and Günter Radden (eds.), 1–46. Berlin: Mouton de Gruyter.

Raible, Wolfgang. 1981. Von der Allgegenwart des Gegensinns (und einiger anderer Relationen). Strategien zur Einordnung semantischer Information. *Zeitschrift für Romanische Philologie* 97: 1–40.

Rettig, Wolfgang. 1981. *Sprachliche Motivation. Zeichenrelationen von Lautform und Bedeutung am Beispiel französischer Lexikoneinheiten*. Frankfurt: Peter Lang.

Roudet, Léonce 1921. Sur la classification psychologique des changements sémantiques. *Journal de psychologie* 18: 676–692.

Saussure, Ferdinand de, 1916. *Cours de linguistique générale*. Paris: Payot.

Schank, Roger C. and Roger P. Abelson, 1977. *Scripts, plans, goals, and understanding*. Hillsdale NJ: Lawrence Erlbaum.

Schwarze, Christoph. ²1995. *Grammatik der italienischen Sprache*. Tübingen: Niemeyer.

Schwarze, Christoph. 2001. *Introduction à la sémantique lexicale*. Tübingen: Narr.

Sharp, Harriet and Beatrice Warren. 1994. The semantics of onomatopoeic words. *Folia Linguistica* 28: 437–447.

Société de Linguistique de Paris (ed.). 2000. *Théories contemporaines du changement sémantique*. Leuven: Peeters.

Swadesh, Morris. 1955. Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics* 21: 121–137.

Swadesh, Morris. 1960 [1972]. What is glottochronology? *The origin and diversification of languages*, Joel Sherzer (ed.), 271–292. London: Routledge & Kegan Paul.

Taylor, John R. ²1995. *Linguistic categorization. Prototypes in linguistic theory*. Oxford: Clarendon.

Ullmann, Stephen. 1962. *Semantics. An introduction to the science of meaning*. Oxford: Blackwell.

Ullmann, Stephen. ²1966. Semantic universals. *Universals of language*, Joseph H. Greenberg (ed.), 217–262. Cambridge MA: The MIT Press.

Ullmann, Stephen. ²1969. *Précis de sémantique française*. Bern: Francke.

Ungerer, Friedrich. 2002. Arbitrarität, Ikonizität und Motivation. *Lexikologie/Lexicology. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen/An international handbook on the nature and structure of words and vocabularies*, D. Alan Cruse; Franz Hundsnurscher; Michael Job and Peter Rolf Lutzeier (eds.), I, 371–380. Berlin: Walter de Gruyter.

Ungerer, Friedrich and Hans-Jörg Schmid. 1996. *An introduction to cognitive linguistics*. London: Longman.

Waltereit, Richard. 1998. *Metonymie und Grammatik. Kontiguitätsphänomene in der französischen Satzsemantik*. Tübingen: Niemeyer.

Wertheimer, Max. 1922/23. Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschungen* 1: 47–58.

Wierzbicka, Anna. 1996. *Semantics, primes and universals*. Oxford: OUP.

Zipf, George Kingsley. 1949. *Human behaviour and the principle of least effort. An Introduction to Human Ecology*. New York: Hafner.

Zöfgen, Ekkehard. 2002. Motiviertheit lexikalischer Einheiten im Französischen, *Handbuch Französisch. Sprache · Literatur · Kultur · Gesellschaft. Für Studium, Lehre, Praxis*, Ingo Kolboom; Thomas Kotschi and Edward Reichel (eds.), 189–193. Berlin: Erich Schmidt.

Authors' addresses:

Romanisches Seminar
Universität Tübingen
Wilhelmstraße 50
72074 Tübingen
phone: ++49.7071.2974296

peter.koch@uni-tuebingen.de
daniela.marzo@uni-tuebingen.de