Viggo Hjørnager Pedersen

REFLECTIONS ON THE TREATMENT OF PREPOSITIONS IN BILINGUAL DICTIONARIES, AND SUGGESTIONS FOR A STATISTICAL APPROACH

Having worked now for some years on the revision of the DANSK-ENGELSK ORDBOG (Vinterberg and Bodelsen), I have come to see dictionary articles as being short, medium, or long. The two first categories have problems of their own; but for both, the arrangement of information within the individual entry is of minor importance; not so with the longer entries of one, two, or more columns: here a suitable arrangement is absolutely imperative in order to ensure speedy access to the information needed by the user.

In a bilingual dictionary like the DANSK-ENGELS ORDBOG (DEO) the long entries almost invariably deal with 'grammatical words' and in what follows I am going to concentrate on the treatment of prepositions.

This is not the place to enter into a detailed theoretical discussion of the concept of 'preposition'. I shall simply take the word in its traditional sense, regarding as prepositions 'particles' (cf. Jespersen 1933:68-69) like English over, to, on, and Danish over, til, på, etc., whenever these words have a complement, as in on the table, or what are you waiting for? However, I shall have to make one or two preliminary observations.

(1) Prepositions may refer to spatial relations (at the table); but almost as frequently they refer to temporal ones (at one o'clock), and very often their concrete sense is lost altogether, and they are reduced to mere relation- or case-markers; they become 'colourless' (cf. Spang-Hanssen 1963, following Brøndal 1943). It is often said about them that they express relations between other words, while being themselves devoid of meaning (cf. Breban 1957). At any rate, it is very difficult to separate their lexical from their grammatical meaning; as we shall see, they are frequently translated in 'grammatical' rather than 'lexical' ways.

(2) Prepositions are normally polysemous, or rather, they enter into such widely different collocations that it may be difficult to find any basic meaning; compare, for example, of in the spirit of romance and all of us, or by in he was sitting by the piano and it was done by me.

As languages are not isomorphic (cf. Catford 1965, Weinreich 1963, Zgusta 1971), it is not surprising that even in the case of closely related languages like Danish and English the polysemous prepositions do not match on a one-to-one basis, not even in the case of etymologically identical prepositions like Danish af and English of. Nor is it possible to establish equivalence for subordinate meanings of roughly corresponding prepositions, though a considerable degree of overlap is normally found.

The most pressing questions for the lexicographer are (1) how to distinguish and indeed to establish the various subordinate

meanings of a given preposition, and (2) in what order to present the translation equivalents. Traditional prepositional theory is of little help here, in spite of the many interesting observations made by writers like Brøndal (1940), Schibsbye (1970), and Spang-Hanssen (1963). Nor are manuals on lexicography like those of Møller (1959), Zgusta (1971), or Al-Kasimi (1977) much better, and while Weinreich pertinently and scathingly criticizes traditional 'anecdotal' lexicography in his 1964 review article, he is not really very helpful himself in his discussion of polysemy in Greenberg's book on universals (Weinreich 1963).

Let us now turn to the dictionary page. If one looks at an entry like E_1 from the DEO, it becomes clear that it consists of an impressive number of subgroups - frequently 10, 20, or even more. The criterion for subdivision is normally - as here - supposed to be semantic difference. But what is the justification for such subdivisions? Here, for instance, sub-entry 4 refers to 'space', 2, 3, and 11 to 'cause' and 'agency', whereas 5, 16, and 19 are simply based on idiomatic expressions. And having asked why the entry on til has 25 sub-senses, whereas af makes do with 19, one may go on to ask the even more difficult question, why the DANSK-FRANSK ORDBOG, which is about the same size, makes do with only 13 sub-entries for af.

E_1 : **1.** *(om materiale, oprindelse)* of (*fx* the house is built of wood ‖ a curtain of silk ‖ a man of humble origin ‖ goods of Danish origin ‖ a car of British make ‖ made of iron); from (*fx* tea made from strawberry leaves); out of (*fx* a cage made out of an old packing-case);
2. *(om den handlende, det virkende, ophavsmanden)* by (*fx* the house was built by an architect ‖ the ship is propelled by a screw ‖ a novel by Dickens);
3. *(ved hjælp af)* by (*fx* he lives by his pen); from (*fx* I see from your letter that ... ‖ it is easily understood from the context);
4. *(væk fra)* off (*fx* he took the lid off the box ‖ he fell off the horse ‖ he helped him off his horse ‖ he washed the soap off his face ‖ he brushed the dust off his coat);
5. *(af ... at være)* for (*fx* highly educated for a peasant ‖ I must admit that for a German he speaks English surprisingly well);
6. *(i henseende til)* of (*fx* gigantic of stature); in (*fx* weak in character), as regards, by (*fx* English by birth ‖ a blacksmith by trade); in the way of (*fx* that is all he has in the way of clothes);
7. *(ved dateer etc)* of (*fx your* letter of May 5);
8. *(om del, blandt)* of (*fx* I want two of these apples ‖ most of them died ‖ one of my friends); out of (*fx* nine out of every ten died ‖ nine times out of ten); in (*fx* one in a thousand);
9. *(stammende fra etc)* from (*fx* I heard it from my sister ‖ I got this watch from my father ‖ I bought it from a man);
10. *(om det karakteriserende, art, egenskab, natur)* of (*fx* a man of my type ‖ a man of enormous strength ‖ a man of high rank ‖ a town of this size ‖ a person of the same name);
11. *(om årsag)* of (*fx* die of hunger); from (*fx* he got a cold from staying out in the rain); with (*fx* he was green with envy ‖ half-dead with fear ‖ the grass is wet with dew ‖ black with soot ‖ tremble with fear ‖ blush with shame); for (*fx* he jumped for joy);
12. *(om motiv)* out of (*fx* he did it out of curiosity (, fear, jealousy, malice, pity, spite)); from (*fx* from politeness, from curiosity);
13. *(om eksistensgrundlag, føde el. penge)* on (*fx* live on vegetables, on a small income ‖ I don't want to live on my wife) *NB sammenlign 3;*
14. *(om samhørighedsforhold, ejendomsforhold etc)* of (*fx* the owner of the dog ‖ the top of the hill ‖ the son of a miner ‖ the discovery of America ‖ a review of the play ‖ the sound of wheels ‖ an admirer of beauty ‖ a portrait of Henry VIII);
15. *(i titel)* of (*fx* King of Denmark ‖ the Governor of California);
16. *(ved nedsættende betegnelser)* of (*fx* that fool of a woman),
17. *det er dumt (klogt, venligt, etc) af dig* that is stupid (wise, kind, etc) of you;
18. *(andre vendinger) hvad er den af?* S what's the big idea? *vi har intet at frygte af ham* we have nothing to fear from him; *det har du rigtig godt af* serves you right; *det var latterligt af en voksen at gøre det* it was ridiculous in an adult to do that; *penge havde han ingen af* he had no money; *det blev der ikke noget af* nothing came of it; *han er en slave af sine lidenskaber* he is a slave to his passions; *han ligger af influenza* he is down with the flu,
19. *mht andre forb med „af", se forb.'s andre led*

The prefaces to large monolingual dictionaries (the OXFORD ENGLISH DICTIONARY, OED, and the ORDBOG OVER DET DANSKE SPROG, ODS) are silent on the subject of criteria for meaning discrimination. Probably there is no criterion apart from intuition, or the authority of older, equally intuitive dictionaries. Thus the system of DEO's sub-senses is heavily influenced by that of the corresponding entries in the ODS.

However, such traditional subdivisions of meaning are unreliable. The OED gives the following information in the entry on love:

> E_2: love (1) That state of feeling with regard to a person which arises from recognition of attractive qualities, from sympathy, or from natural ties, and manifests itself in warm affection and attachment ...
>
> (4) That feeling of attachment which is based upon difference of sex; the affection between lover and sweetheart.

But it is very difficult if not impossible really to distinguish between these senses: "I love my grandmother" presumably would be (1), but what about "I love my cousin"? The same obviously applies to the semantic subdivisions of dictionary entries on prepositions. Thus a German-English dictionary under bei has a sense 1, defined as "proximity in space", and under 1d the example "bei wem hast du Stunde"? But under 2, undefined, we have "bei ihr ist alles möglich". It would have been interesting to see where the editors would have placed "bei mir bist du schön"; but the answer is of course that there is no sharp dividing line between senses 1 and 2, which rather blur into each other (cf. the paper on meaning discrimination in monolingual dictionaries by Stock in Part I of this volume).

The approach of the studies which I am going to discuss here is different, in that they are all mainly concerned with the relative frequency of the translation equivalents of Danish prepositions. My material consists of five such studies, dealing with six prepositions in all, together with three other investigations which do not fit in with the rest, but whose findings are nevertheless analogous.[1] Examples have been taken from modern Danish texts, fictional and non-fictional, translated into English by professional translators. The papers were written by students at the Copenhagen School of Economics and Business Administration as part of the requirements for the intermediate EA (= BA) level. All except the papers on i and om were written under my direction, but a standardized method of investigation was only gradually evolved.

I have felt for some time that it must be possible to pool the results of these investigations, and to draw some general conclusions as to the adequacy of the entries in DEO. Looking at Figs. 1 and 2, one is struck above all by the number of equivalents - for ved no less than 40. Next, one notes the prominent position of paraphrase and deletion - which, taken together, constitute the second most frequent category of all. Fig. 3 tabulates the results of the main projects. Especially interesting

	RR	TT	OO	AS	AB	AR	PE	TOTAL ABS.	TOTAL %
OF	43	72	87	75	22	18	26	343	47,5
OMSKRIVNINGER	42	15	10	29	8	6	13	123	17,0
BY	5	12	21	19	1	2	3	63	8,7
IKKE OVERSAT	29	6	7		1	4	1	48	6,7
FASTE PRÆP.FORB.	2	7	5	12			5	31	4,3
OUT OF	5	13				1		19	2,6
FROM	5	5	5					15	2,1
ING-FORM		2	5	2	1	1	1	12	1,7
FASTE FORB.	7	1	2					10	1,4
FOR		2	1	2		1	1	7	1,0
IN			1	2	1		3	7	1,0
ON			2	5				7	1,0
WITH	4	1					1	6	0,8
BESIDE	1	3						4	0,6
OFF	2							2	0,3
NEXT TO	1	1						2	0,3
DUE TO				2				2	0,3
THROUGH				2				2	0,3
S-GENITIV		2						2	0,3
POSS.PRON.		1						1	0,1
AT		1						1	0,1
NEAR	1							1	0,1
ALONGSIDE	1							1	0,1
ADV.									
NOW & THEN/AGAIN	2	3						5	0,7
SUCH AS			1				1	2	0,3
OFF		1						1	0,1
AS			1					1	0,1
LIKE			1					1	0,1
ONCE IN A WHILE		1						1	0,1
ACCORDINGLY		1						1	0,1
NECESSARILY			1					1	0,1
	150	150	150	150	34	33	55	722	99,9%

Fig. 1: English translation equivalents of the Danish preposition af (based on 722 examples)

	FF %	HU %	LØ %	MO %	TOT. ABS.	GNS %	TOT. ABS.	GNS %
1. at	14	19	13	22	68	17,-	93	18,6
2. by	28	15	18	13	74	18,5	85	17,-
3. udvidet tid	12	11	3	1	27	7,-	36	7,2
4. ing-form	-	3	2	-	5	1,25	8	1,6
5. about	3	3	4	-	10	2,25	11	2,2
6. about to	2	5	1	-	8	2,-	13	2,6
7. against	-	1	-	-	1	0,25	2	0,4
8. almost	1	-	4	-	5	1,25	6	1,2
9. along(side)	1	-	1	2	4	1,-	4	0,8
10. behind	-	2	1	-	3	0,75	7	1,4
11. below	-	-	1	-	1	0,25	1	0,2
12. beside	1	10	6	1	18	4,5	22	4,4
13. close to	-	-	-	1	1	0,25	1	0,2
14. during	-	-	-	7	7	1,75	7	1,4
15. for	-	1	1	3	5	1,25	5	1,-
16. in	7	5	2	10	24	6,-	25	5,-
17. in front of	2	-	1	-	3	0,75	3	0,6
18. in relation to	-	-	-	1	1	0,25	1	0,2
19. in the case of	-	-	-	1	1	0,25	1	0,2
20. in the course of	-	-	-	1	1	0,25	1	0,2
21. in the process of	-	-	-	1	1	0,25	2	0,4
22. near(to)	1	2	1	-	4	1,-	11	2,2
23. nearest	-	1	-	-	1	0,25	2	0,4
24. nearly	-	1	3	-	4	1,-	5	1,-
25. next (to)	-	5	2	-	7	1,75	7	1,4
26. of	2	1	3	2	8	2,-	9	1,8
27. on	7	4	6	9	22	5,5	24	4,8
28. on top of	-	-	-	-	-	-	1	0,2
29. on the point of	1	-	-	-	1	0,25	2	0,4
30. on the verge of	-	-	-	-	-	-	1	0,2
31. over to	-	2	2	-	4	1,-	4	0,8
32. out to	-	1	-	-	1	0,25	1	0,2
33. round	-	-	1	-	1	0,25	1	0,2
34. through	-	-	2	1	3	0,75	3	0,6
35. to	3	-	-	-	3	0,75	3	0,6
36. to + inf.	3	-	-	-	3	0,75	6	1,2
37. towards	-	-	2	-	2	0,5	2	0,4
38. when	1	-	-	3	4	1,-	5	1,-
39. with	5	-	9	5	19	4,75	23	4,6
40. omskrivn.	10	8	11	16	45	11,25	56	11,2
TOTAL	100	100	100	100	400	100	500	100

Fig. 2: Translation equivalents of the Danish preposition ved, based on 500 examples (absolute figures and average percentages in the two right-hand columns)

are Columns 2, 5, 8, and 9. Looking at the averages, it will be seen that the average number of equivalents per preposition is 31.3, but that nevertheless paraphrase and deletion make up no less than 21.3%, or more than one fifth of the total number of examples.

This category obviously calls for further investigation, but it certainly serves to remind us that we do not translate words as much as phrases and sentences, a fact which is insufficiently mirrored by the DEO and, I think, a good many other bilingual dictionaries.

	1	2	3	4	5	6	7	8	9
Danish prepositions	Number of examples	Number of equivs.	Paraphrase (%)	Deletion (%)	Categories 3+4 (%)	Most frequent equiv. (%)	2nd most frequent equiv. (%)	Sum of 5 most frequent equivs. (%)	Sum of remaining equivs. (100% - columns 8+5)
af	722	23	17	6.7	23.7	47.5	8.7	65.2	9.1
over	481	37	11.2	4.1	15.3	22.9	9.4	43.2	41.5
på	668	38	24	6	30	25	9	55	15
til	800	24	19.3	-*	19.3	45.8	16.1	69.8	10.1
under	239	26	14.7	3.3	18	37.6	7.1	55.5	26.5
ved	500	40	21.4	-*	21.4	18.4	17	49.8	28.8
average	568	31.3	-*	-*	21.3	32.9	11.2	56.4	21.8

equiv./equivs. = equivalent/equivalents

Fig. 3

* The papers on til and ved include the category of deletion in that of paraphrase. Consequently, no separate averages have been worked out for these categories.

One important source of deletion is immediately obvious: in Danish (unlike English), prepositions may govern that-clauses, and the use of prepositions immediately before subordinate clauses is more widespread in Danish than in English. According to the traditional Danish grammars and dictionaries, a frequent choice for a translator faced with a Danish preposition + at-sætning (that-clause) is to add a prop-word after the preposition: "Han var irriteret over, at der blev talt amerikansk" (He was annoyed by the fact that American was spoken) or "De skal nok sørge for, at de ikke går glip af noget" (They'll see to it that they don't miss out on anything). However, the paper by Albeck shows that this construction is comparatively rare, and that a much more frequent solution is to leave out the preposition: "... for at prale med, at du er bedre til at tale engelsk" (To boast that you're better at speaking English) or "Han var ikke klar over, at hun var kommet" (He was unaware that she had arrived).

As for paraphrase, it has been demonstrated, particularly in Krog's paper on ved, that prepositional phrases may be replaced by verbal constructions, such as -ing-constructions, subordinate clauses, or continuous tense: ved indlæggelsen (when he had come into hospital); ved stiftelsen af et A/S (when forming a company); ved hjælp af skrivebordets kant (using the edge of the desk); han var ved at iføre sig sin frakke (he was just putting on his coat).

The selection of the five most frequent direct equivalents (i.e. exclusive of paraphrase) is more or less arbitrary. Some of the papers had chosen this number because it fitted their material, and it has been retained in the comparative survey. The idea was to see what degree of coverage could be obtained by a relatively small number of equivalents, something which would certainly be of interest for machine translation, as discussed by Lenstrup. As the table in Fig. 3 demonstrates, the five most frequent equivalents make up about 50%, frequently more. The remainder are mainly low-frequency items, as appears from column 9, which shows the percentual coverage of the sum of all equivalents other than the top five, plus deletion and paraphrase. As will be evident, the rarest equivalents do not add up to much. Thus the bottom 14 equivalents of af together make up no more than 6.3% of the total number of translations. This points to the conclusion that whereas it would not be possible to build a dictionary entry around the high-frequency equivalents alone, there is every reason to give them prominent treatment in the entries.

Let me quote at this point the conclusion of one of the student researchers, Persson:

> As it has only been possible to find examples of about half the categories using of included in the DEO entry on af, I think that the remainder must be of such low frequency that they hardly need separate treatment. They could be lumped together under one head, and just be mentioned with one example each. I also think that the entry on af should deal with translation equivalents according to their relative frequency, and as of is the most frequent equivalent, it would seem natural to deal with this word first, grouping its semantic subcategories according to frequency.

This would be a radical solution indeed. We are not going to adopt it in the next edition of DEO, largely because it will not be possible to investigate the translation of all high-frequency items. Still, changes in the entries for which there is statistical material to draw on will certainly be made.

There is insufficient space to embark on a detailed discussion of all 6 prepositions, but in what follows we will look more closely at af. The DEO entry on af is subdivided according to a semantic system close to that of ODS; but DEO is a bilingual dictionary, and the same translation equivalent frequently appears in more than one semantic subgroup. Proposed equivalents are: of, out of, from, off, for, in, with, on, to, plus various periphrastic translations. Persson has the same equivalents, - to, but + through, at, near. In other words, there is a fair degree of correspondence; but it is not clear from DEO how absolutely dominant of is as a translation (47.5%), and how far the order of the DEO subdivisions is from being based on considerations of frequency of occurrence (cf. Fig. 4, from Persson, p.14).

		ABS.	%
14	possession, connection	220	64
8	part of, among	80	23
6	as regards	14	4
1	material, origin	12	3
10	nature, type, property, what characterizes	10	3
11	cause	6	2
	not compatible with the DEO system	1	1
		343	100%

Fig. 4: Labels from the DEO entry on af (translated). The figures on the left indicate the place of the various sub-senses in the DEO entry, those on the right their frequency of occurrence. Only sub-senses to which of is a possible equivalent have been considered; even so it is remarkable that only six categories have been found, and that the most frequent category is listed as number 14 in the DEO.

It is not possible - at least for the present - totally to rewrite entries like that on af. But it is possible to reduce the large number of sub-senses, and to rearrange the rest so as to correspond more closely to an order based on frequency of occurrence, both of source-language sense and of translation equivalent.

Referring back to entry E_1, I propose to rearrange it somewhat

as follows: Sub-senses 14, 8, 1, 10, 11, 6, 7, 2, 3, 4, 9, 12, 13 will form sections 1-13 of the new entry. Senses 15, 16, 5, 17 will be run together into a new section 14, the old senses 18 and 19 becoming 15 and 16 in the revised entry. Obviously a number of other changes are called for, and generally speaking the revised articles must pay more attention to periphrastic translations than the old ones. But the principle of having, as far as possible, the most frequent sub-senses first in a long entry does seem to deserve more attention than it has had so far.

In conclusion, I should like to extend my thanks to the dedicated students who have provided the figures which form the basis of this report, and of the revision of the relevant dictionary entries. It is to be hoped that more such studies will be forthcoming, so that in time there will be statistical material on which to base revised versions of all the DEO entries on 'grammatical words'.

Note

1 The following unpublished undergraduate dissertations for the EA degree at the Handelshøjskolen i København have been consulted:
Albeck, U. English Translation Equivalents of Danish Prepositions with Subordinate Clause Complement
Eriksen, L. English Translation Equivalents of Danish 'til'
Krog, J. English Equivalents of Danish 'ved' Arranged According to Frequency
Lenstrup, R. Mechanical Translation of the Word 'om' from Danish to English
Molich, I. English Translation Equivalents of the Danish Prepositions 'over/under'. A Contrastive Analysis
Outzen, K. English Equivalents of the Danish Preposition 'på' Arranged According to Frequency
Persson, B. English Equivalents of the Danish Preposition 'af'
Waterhouse, U. English Translations of the Local/Geographical Senses of the Danish Preposition 'i'

References

Al-Kasimi, A. (1977) Linguistics and Bilingual Dictionaries. Leiden: Brill
Breban, V. (1957) "Les prépositions dans les dictionnaires unilingues" in Mélanges Linguistiques. Bucharest: Académie de la République Roumaine
Brøndal, V. (1940/50) Praepositionernes theori (French translation Théories des prépositions). København: Lunos
Brøndal, V. (1943) "L'originalité des prépositions du français moderne" in Essais de linguistique générale (by Brøndal). København: Munksgaard
Catford, J. (1965) A Linguistic Theory of Translation. London: Oxford U.P.
Jespersen, O. (1933) Essentials of English Grammar. London: Allen and Unwin
Møller, K. (1959) Leksikologi og leksikografi. København: Schultz
Schibsbye, K. (1965/70) A Modern English Grammar. Oxford

Spang-Hanssen, E. (1963) Les prépositions incolores du français moderne. København: Gads Forlag

Stock, P. (1984) "Polysemy" in this volume

Weinreich, U. (1963) "On the semantic structure of language" in Universals of Language ed. by J.H. Greenberg. Cambridge, Massachusetts: M.I.T. Press

Weinreich, U. (1964) "WEBSTER'S THIRD: a critique of its semantics" International Journal of American Linguistics 30: 405-409

Zgusta, L. (1971) Manual of Lexicography. The Hague: Mouton