

**Grammatical Information in the New “Bi-directional” Polish - Dutch Dictionary:
Some Problems and a Few (Im)practical Solutions**

René Genis, University of Amsterdam (published: 2002. [in:] *Lexicography and Language Policy in South-Slavic Languages after 1989* (= *Die Welt der Slaven* 14), [ed.] Radovan Lučić. München.)

1. introduction

It is probably impossible to produce a truly bi-directional bilingual dictionary. In order to describe both languages (n and p) accurately and to accommodate both sets of users (N and P) ideally, it would be necessary to make four parts; two in both directions for each of the two groups of users. The respective natives of each language would not be bothered then by metalingual information (*n* and *p*) concerning their own language.

Schematically this non-bi-directional situation would be:

	source language	target language		meta-language	targeted user
part 1	n	p	}	<i>n</i>	N
part 2	p	n			
part 3	n	p	}	<i>p</i>	P
part 4	p	n			

table 1; *ideal situation*

However, where (a.o. commercial!) interest in a dictionary for two particular languages is relatively small, it pays to come to certain compromises so that only two parts need be compiled.

	source language	target language	metalanguage	targeted user
part 1	n	p	<i>n</i> (on p for N)	} N + P
part 2	p	n	<i>p</i> (on n for P)	

table 2; *bi-directional situation*

Obviously, the above table shows the situation in a necessarily oversimplified way; matters metalingual are never as straightforward as that. For one, there is often a possibility of solving the bilingual needs at the metalingual level by using means that are comprehensible for both N and P. For instance, in our dictionary it has occasionally been possible to utilise grammatical information (about p for N; where p is the source language) concerning the government of verbs to differentiate a lexeme semantically (about n for P; where p is the source language). Table 2 represents the intention with which we are currently making a Polish - Dutch and Dutch - Polish dictionary at the universities of Warsaw, Wrocław and Amsterdam.¹

It goes without saying that all lexical information should be presented in such a way, that it is readily available for the group of users that need it and yet unobtrusive for those that do not. However, the question — in the case of our own dictionary as yet largely unanswered — as to how exactly certain information is to be presented was to a large degree left to the typographers of our publisher, Van Dale Lexicografie bv, Utrecht / Antwerp.

The present paper will not go into the typographical side of matters further than absolutely necessary. We are concerned here primarily with the question whether, and if so, which grammatical data should be included in our part of the newly to be developed Polish - Dutch dictionary. There is, of course, a case for not including any such information at all. Users should simply reach for a grammar (or perhaps a good monolingual dictionary) when the need for inflectional information concerning a particular lexeme arises. Unfortunately, the situation is such, that there are no accessible grammars readily available (in Dutch) that would deal with all the necessary grammar for Polish in such a way, that would make this a real option for the Dutch speaking user. Monolingual dictionaries are often (too) cumbersome in use and require one to do at least one, sometimes even more bouts of looking up in order to find the coveted information.

¹ The author is only involved in the making of the part Polish-Dutch and not Dutch-Polish of what is eventually to become a two-part dictionary. The matters under discussion in this paper concern the forthcoming part Polish - Dutch.

Quite apart from this, there is a certain tradition of providing grammatical information for Polish in bilingual dictionaries, where Polish is the source language. This usually concerns paradigmatic particulars. When it comes to ease of use, one could argue that it would be better to provide such information not only in the Polish - Dutch, but certainly also in the Dutch - Polish volume. In the latter a Dutch user would look up a word to be translated into Polish and get the full paradigm with the translation provided. Many users also use dictionaries to find out or check grammatical information about lexemes that are already in fact known to them (at least to some degree). Obviously it would be very user-friendly if grammatical information could be found in the same volume one uses to find one's translations.

Moreover, when one looks at Polish as opposed to Dutch, the former has a rather complicated and confusing inflectional system. Table 3 demonstrates that it can at times be quite confusing for the less initiated to grasp. It includes all the case endings for substantives throughout the various paradigms.²

	singular	plural
Nominative	-∅-a -o -e -i -ę	-i -y -a -e -owie
Genitive	-a -u -y -i	-i -y -∅-ów
Dative	-owi -u -'e -i -y	-om
Accusative	-∅-ą -o -e -i -ę	-i -y -∅-ów -a -e
Instrumental	-em -ą	-ami
Locative	-e -i -y -u	-ach -"ech
Vocative	-o -u -i -'e -"e	-i -y -a -e -owie

table 3; case endings

You will notice that certain endings appear several times, others three or four times and some only once. The prize for most frequently occurring inflectional ending goes to *-i*, which is present in no less than 10 instances! Imagine a not-so-well-informed Dutchman

² In table 3 -'e, -'i indicate endings with "softening" such as *lampa : lampie*, while -"e refers to such alternations as *ojciec : ojcie*, and -"ech to those like *Niemcy : Niemczech*.

(or Dutch speaking Belgian for that matter) trying to find out which case or number he is dealing with when a word ending in *-i* crops up in an example provided. And let's not forget that the majority of users will actually be not-so-well-informed. Although matters are not quite as complicated as they might seem from the above—a fair number of lexeme-classes do not have questionable or confusing paradigms—this nevertheless makes a case for providing accurate and complete paradigmatic data with each lexeme. This, at least, is the view of our editorial team.

Thus, having decided that inflectional data needed to be given for Polish, we had to decide exactly which data we would include and on which level. In this respect we discern three levels:

- A. headword, entry;
concerns all lexemes with a paradigm;
includes word-class label, paradigm
- B. government;
concerns verbs and prepositions;
- C. phraseology, examples.

2. Some existing types of dictionary compared

I want to have a look at certain strategies for the inclusion of grammatical information in particular, in some of the existing dictionaries. There appear to be three types of dictionaries:

1. dictionaries with only limited grammatical information
2. dictionaries which state the main exceptions as well as the more unexpected forms, but which do not provide complete paradigmatic data
3. dictionaries which (claim to) provide full paradigmatic data

For each of these types I have selected one or two dictionaries from which to draw examples from:

- (1) Martens Nico, Morciniec Elke, *Klein Nederlands - Pools en Pools - Nederlands woordenboek*, Wiedza powszechna, Warschau 1977
- (2a) Stanisławski Jan, *Wielki słownik polsko - angielski*, Warszawa 1983
- (2b) Fisiak Jan, *Słownik polsko - angielski*, Collins - BGW, Warszawa 1996
- (3a) Van den Baar A.H., *Russisch - Nederlands Woordenboek*, Aula (210), Antwerpen - Utrecht 1965
Van den Baar A.H., *Nederlands - Russisch Woordenboek*, Antwerpen - Utrecht 1994
- (3b) Dunaj Bogusław, *Słownik współczesnego języka polskiego*, Wilga, Kraków 1996

You will notice that (1) through (3a) are bilingual dictionaries, while (3b) is monolingual, Polish. I will look at these per level A, B and C. The dictionaries under (3a) are not discussed in full; Van den Baar [1994] is looked at in the discussion on paradigmatic data at headword level, while Van den Baar [1965] comes up for the discussion on government.

Of course the dictionaries under scrutiny were not meant to be bi-directional, so they could only serve as sources of inspiration for the design of our own dictionary, not as models. (2b) is slightly hybrid as its novel approach gives rather more information than dictionaries of type 1, but certainly not as much as type 3 dictionaries.

Following this I will discuss some of the solutions we came up with, working within the framework of the Van Dale - style of dictionary, as we had to do.

2.1. Level A - headword

We found various solutions and choices as to the inclusion of paradigmatic particulars at headword level.

The limited scope and size of the only Dutch - Polish and Polish - Dutch dictionary in existence by Martens & Morciniec (1) allows its authors to give minimal paradigmatic information only. The authors' main concern are the alternances of masculine nominal stems. No such data is provided for feminine or for neuter nouns. Thus, complicated alternances like *jazda* : *jeździe* are supposed to be found in other sources, while more obvious ones such as *groszek* : *groszku* and *gród* : *grodu* are included solely for the fact that they are masculine. Martens and Morciniec chose to present these alternances for masculine nouns in the shape of the genitive singular. The choice of that particular case is a good one as it is one of the most problematic cases as far as choice of ending is concerned in Polish: masculine (non animate) stems take either *-a* or *-u* and the rules are rather convoluted. In addition to stems that alternate and take one of the two possible endings for this case, the authors indicate the ending for the genitive singular of several other masculine nouns, although there does not seem to be a clear line in their policy to do so though. For instance, the first meaning of *przewodnik* is 'guide [person]', which has a genitive singular in *-a*. The second meaning is 'guide [booklet]', which takes *-u*. The ending for the first meaning might be considered obvious, as one of the main guidelines for this case states that living beings take *-a*. The ending *-u* for the second meaning can be considered fairly unexpected (especially as most words ending in *-ik* take *-a*). Neither form is included. This is surprising, as the authors promise in their grammatical introduction to "state any genitive singular about which there could be confusion" (p. 447-2). A similar example is *gruchot* for which we find three meanings. The first of these, (i) 'rattle' is accompanied by the genitive ending *-u*. The following two meanings (ii) 'old rubbish' and (iii) 'old person' are left unmarked as to genitive case. In fact both (ii) and (iii) take *-a*, and in the view of our team it would have been nice to be given these forms, as there could be uncertainty for the Dutch user. As it is, this particular lemma is fairly confusing; by stating *-u* for meaning (i), it is unclear whether meanings (ii) and (iii) should get *-u* or *-a*. Further inconsistent examples can be found among loan words such as *grotmaszt* 'main mast' for which is given the possibly not so surprising (borrowing, hard stem, large object) genitive ending *-u*, while the rather more frequent *komputer* 'computer' goes without its probably unexpected *-a*. Interestingly, none of the other obscure case-endings for masculine nouns are elucidated. Neither are we enlightened about the choice between *-owie* and *-i* for the nominative plural, nor about *-ów* or *-i* for the genitive plural. Perhaps most surprising of all, although admittedly consistent, is the

complete lack of information about rare and unexpected case endings such as the dative singular for words like *kot*, *pies*, *chłopiec*, *ojciec*, *pan*, *świat*, which all take *-u* instead of the regular *-owi*.

No information (apart from a separate list of irregular verbs with their respective main forms on pp. 464, 465) on the conjugation of verbs or the declension of any other part of speech is included. Furthermore, the sporadic examples and phraseology provided are not very helpful in this respect.

All in all, we can conclude that this particular dictionary raises a corner of the paradigmatic veil in some cases but leaves the user with many questions and uncertainties in others.

Both type 2. dictionaries under scrutiny include partial paradigmatic information. The limited space allowed for this paper forces me to restrict myself to the treatment of substantives. We find that Fisiak (2b) gives the genitive singular and the nominative plural for all nouns masculine, feminine and neuter where such forms exist as a standard accessory. Fisiak adds to this any “further irregularities in declension” (sic; p. III). In fact, Fisiak is a lot kinder to his users than he himself claims to be; some quite regularly formed cases are marked consistently. These include the instrumental singular for neuter and masculine stems in *-k/g(o)* such as *niewiniątko* : *niewiniątkiem* ‘innocent’, *pstrąg* : *pstrągiem* ‘trout’ and *grozdek* : *grozkiem* ‘green peas’. Apparently the author thought these forms worthy of the epithet “irregular” as they have an inserted *-i-*. Not surprisingly, Fisiak also adds the complicated, very frequent, but certainly not “irregular” forms for the neuter and masculine locative singular as well as the feminine dative singular. Thus he gives *kot* : *kocie* ‘cat’. However, I have not been able to find any mention of the rare masculine dative singular in *-u* for *kot* : *kotu* ‘cat’ nor for the equally frequent lexeme *pies* : *psu* ‘dog’, although most of the others of this class are appropriately marked. The probably equally rare masculine accusative singular in *-a* for inanimate nouns such as *trup* : *trupa* ‘dead body, corpse’, *mazur* : *mazura* ‘mazurka’ and *walc* : *walca* ‘waltz’ is not specified.

Stanisławski (2a) also treats masculine, feminine and neuter nouns with equal respect and “sporadic occurrences” and “peculiar” and “exceptional” (sic; pp 812 and following) case endings are usually mentioned. So far, I have noticed but one oversight which concerns the rare masculine accusative singular in *-a*. Although given as examples of this

“exceptional” formation in the part entitled *Elements of Polish Grammar* at the end of the second volume (pp. 907 and following), at headword level the lexemes *mazur* ‘mazurka’ and *trup* ‘dead body, corpse’ are not marked for this case. *Walc* ‘waltz’ is not given as an example in the aforementioned compendium but is marked for this case in its entry.

In addition to the above, Stanisławski treats the masculine genitive singular by way of a ‘default’ arrangement: when no information is given one should assume *-a* for this case, otherwise the *-u* is stated.

Both type 2. dictionaries do in fact treat stem alternances of nouns very well; they are virtually always mentioned.

It would seem that whenever the decision is made to include only partial paradigmatic data at headword level, inconsistencies and omissions arise which are confusing to the user and which often leaves him / her in the lurch. Moreover, the makers of these two dictionaries have made the choice that for common, unexceptional case forms the user is still compelled to use a grammar.

Dictionaries that include full paradigmatic information can be divided into two general types, which are exemplified here by (3a) and (3b).

Dunaj (3b) is the most recent representative of this, the most common approach for providing paradigmatic data in Polish monolingual dictionaries that was available to us at the outset of the work on our dictionary. All major recent dictionaries employ this method.³ It consists of a grammatical compendium at the front of the dictionary, which presents numbered tables of full paradigmatic prototypes. A prototype number referring to this compendium accompanies each headword. Moreover, where deviant case forms occur they are written out at headword level. In a way, this system represents the best of both worlds for not only does it enable the user to find full paradigmatic information, it also signals special case forms at headword level. In order to keep the amount of prototypes as restricted as possible the author of this dictionary had to marry certain morfo-fonological alternants within one prototype. For instance, the headword *wilga* ‘golden oriole’ has a number that refers to the paradigm of *sroka* ‘magpie’ at the beginning of the volume. As the dative / locative singular for these two lexemes do not

³ I have not yet been able to study the most recent important dictionary: Bańko, Mirosław, *Inny słownik języka polskiego*, Wydawnictwo Naukowe PWN, Warszawa 2000

concur: *sroka* : *sroce* (alternance k : c) ~ *wilga* : *wildze* (alternance g : dz) Dunaj is forced to enter this form for *wilga* at headword level. In practice one finds that in about 85% of entries this type of ‘additional’ paradigmatic information takes up at least the first half of the top line of the lemma, while some 10% spend more than one line.

In principle it is not the task of a bilingual dictionary to inform native Poles about their own language. This is where they differ from monolingual dictionaries. In view of this, one could argue that the space taken up by paradigmatic data at headword level in bilingual dictionaries should be as little as absolutely possible, so as not to bother the native Polish user. In Van den Baar [1994] (3a) we find an approach that accommodates such requirements. It is similar to Dunaj in that it uses numbers to refer to a grammatical compendium. Although in this dictionary it is the target language, Russian, which is described — hence the numbers are to be found not with the headword but with the translation — it is easy to imagine this situation reversed.⁴ The main difference with dictionaries like Dunaj’s is the way in which the compendium is compiled. It consists of many more types than Dunaj’s, such that for example *wilga* and *sroka* would have had separate numbers, had this been a Polish dictionary. In fact, the number of types runs up to 461 for all parts of speech together. In comparison with Dunaj, who has about 76 prototypes, this seems rather a lot. Is it right to say that these dictionaries provide paradigmatic information at headword level though?

2.2 Level B - government

One could argue that this is really headword level information albeit specific for verbs and prepositions. As it is quite different in character to the paradigmatic matters discussed above I will look at this type of information separately. Treatments differ for verbs as opposed to prepositions. I will look at verbs first.

Type 1. dictionaries usually do not go into this matter in a structural way. In Martens & Morciniec (1) there are simple labels to indicate transitivity but no other means appear to have been used. Examples and phraseology are not designed to elucidate matters further.

⁴ A better example of this practice might have been: Kopeckij L.V., Filipec J., Leska O., *Čessko-russkij slovar*, Čechoslovackaja akademija nauk - Institut inostrannyh jazykov, Moskva 1973 and possibly earlier Czech - Russian dictionaries as well. I was not aware of this fact at the time of preparing this paper. I am indebted to prof. Jiří Marvan of the University of Prague / Pilsen for pointing this out to me.

The type 2. dictionaries in existence for Polish usually adhere to what might be called the traditional method in use for Slavic languages. It employs the words *ktoś* ‘somebody’ and *coś* ‘something’ in the appropriate case for the verb they describe, e.g.

<i>uczyć</i>	<i>kogo czego</i>	‘to teach sb sth’
<i>żałować</i>	<i>czego</i>	‘to regret sth’
<i>ufać</i>	<i>komu, czemu</i>	‘to trust sb, to trust in sth’
<i>walczyć</i>	<i>z kimś o coś</i>	‘to fight with sb for sth’
<i>władać</i>	<i>czymś</i>	‘to rule / wield sth’

Stanisławski (2a) faithfully adheres to these formulae. Fisiak (2b) does not always give such a formulaic elucidation but rather opts for an approach using selected examples. His entry for *uczyć* gives and translates the example *uczyć (kogoś) matematyki/polskiego* ‘to teach (sb) mathematics/Polish’, but does not present the formulaic basis to this example. By the same token he has several examples to illustrate the usage of *walczyć*: *walczyć ze snem* ‘to fight off sleep’; *walczyć ze śmiercią* ‘to be fighting with death’; *walczyć o / za coś* ‘to fight for sth’; *walczyć z czymś* ‘to fight against sth’. Again the basic formula is lacking. Both *uczyć* and *władać* are labeled for transitivity in Fisiak. The label for the latter is extended with lexicographic information about the obligatory instrumental case. It looks like this: *vt +instr.*

It is difficult to say which of the methods adhered to in these two dictionaries is the easier to comprehend for a non-Polish native.

A quite elegant as well as space saving solution is offered by Van den Baar [1965] (3a). Here we see the use of numbers in square brackets to indicate the case(s) to be used. Thus we find e.g.: *обучать* [4], [3] ‘to teach sb sth’. This has the added advantage that cases are actually named, which makes following examples a lot clearer.

Monolingual dictionaries such as Dunaj (3b) never give formulae of the kind employed by Stanisławski (2a). They usually illustrate matters through examples of usage, without actually naming the cases used.

When it comes to the treatment of government of prepositions I can be brief.

Martens & Morciniec (1) does not specify case in a metalingual way at all. The examples are not ordered per case.

Stanisławski (2a) is essentially a dictionary meant for Polish users. It states the case used for certain meanings of the prepositions in Polish. The lemma is organised according to case.

Fisiak (2b) employs labels indicating case and organises the article according to case. As these labels are bilingual in essence, they can be comprehended easily by natives of both languages.

All the recent type 3. dictionaries use a method to describe government of prepositions similar to the one employed by Stanisławski (2a); no formulae are given but examples abound.

2.3. Level C - phraseology, examples

None of the types of dictionary elucidates case used in examples and phraseology in a systematic way.

3. Choices, solutions for Van Dale's Polish - Dutch

The following table shows the basis for our choices according to type of dictionary and level of information:

level	method as in dictionary type
A - headword	(3a)
B - government	(2a) + additions, (2b)
C - example, phraseology	innovations

table 4; *Van Dale Polish - Dutch*

Partly in view of the above conclusions about type 1 and type 2 dictionaries our team decided to go for an approach that would make it possible for the user to find out full

paradigmatic details within the dictionary itself. We have opted for an approach as in Van den Baar [1996]. This means that no actual case forms are given at headword level at all. Instead, every lexeme with a paradigm is given a number referring to a very extensive compendium, that covers absolutely all morfo-fonological types. These tables were devised by dr hab. Włodzimierz Gruszczyński of the University of Warsaw. We consider this the most convenient way for both groups of users, as on the one hand it provides all possible inflected forms but on the other it is not obtrusive for such users that do not need this information. A disadvantage of this method is the lack of “warning” at headword level for the uninitiated. Surprisingly, the amount of morfo-fonological types established far exceeds the number of 1000, which makes the compendium quite bulky. Hopefully our codes will eventually be used in the Dutch - Polish part of our dictionary as well. I have not mentioned the paradigmatic complications concerning verbs and other parts of speech that we decided to elucidate for our Dutch-speaking users. The treatment for these will be exactly as those for the substantives.

Labels will indicate government for prepositions in our dictionary. The articles will be arranged according to case governed, and there will be a lot of illustrative examples. Labels and examples will also abound to illustrate the government of verbs⁵, but we will add formulae of the type described for Stanisławski (2a) as well. We had to modify the appearance of these to make them absolutely unequivocal for the Dutch user. As these formulae now had to specify the use of these verbs for the non-Polish user (as opposed to Stanisławski’s) we had to devise a way to avoid the risk of confusion about the case factually indicated. For a Dutch user Stanisławski’s description *uczyć kogoś czegoś* ‘to teach sb sth’ might lead to the erroneous production of **Chłopiec uczył dziewczyny polskiego*, i.e. with a genitive rather than the correct accusative.⁶ We found that over 40% of our second year students (who have a very reasonable knowledge of Polish) would confess to being unsure of the exact case to use and found the information provided by Stanisławski “inconclusive”. This meant that although Stanisławski gives a hint of what is going on with the government of this particular verb, there was still a clear need to look

⁵ We have four labels for verbs: 1. to indicate intransitivity, 2. to indicate transitivity with an accusative object, 3. to indicate transitivity with non-accusative objects, 4. to indicate reflexivity.

⁶ Stanisławski adheres to the Polish tradition of labeling as transitive, verbs with objects in other cases than the accusative. This does not lessen the confusion for the non-Polish user.

into the matter further. The main problem with Stanisławski's versions lies in the fact that various case forms of *ktoś* and *coś*⁷ are in fact homonymous. In our dictionary the homonymic forms will be marked with a superscript small capital indicating the case intended. The full paradigms for these words as used in our dictionary will be:

genitive	<i>kogo</i> ^G	<i>czego</i>
dative	<i>komu</i>	<i>czemu</i>
accusative	<i>kogo</i> ^A	<i>co</i>
instrumental	<i>kim</i> ^I	<i>czym</i> ^I
locative	<i>kim</i> ^L	<i>czym</i> ^L

In our dictionary the above example will show as *uczyć kogo^A czegoś*, which will give the Dutch user a fair chance to produce the correct *Chłopiec uczył dziewczynę polskiego*.

It is our intention to include case markings at the level of the examples and phraseology. Often phrases have a “fixed” combined with a “variable” part such as in *dobry z matematyki*. The fixed combination of *dobry z* (+ genitive) is linked to the variable *matematyka*. Instead of *matematyka* we might want to use *biologia*, *czytanie*, *geografia* etc. In order to do so we need to know that the variable needs to be put in the genitive case (rather than the instrumental, which might also be possible with *z*). In our dictionary we want to remove any doubt by marking the case for variable elements in examples, thus: *dobry z matematyki^G*.

Hopefully the choices we have made for our Polish - Dutch dictionary will be useful for both Polish and Dutch users. Still we do not expect that we will have produced the ideal bi-directional Polish - Dutch dictionary.

⁷ we chose to use versions without the final *-ś*, as this is usually the way Poles memorise the use of case.

Bibliography

Van den Baar A.H., *Russisch - Nederlands Woordenboek*, Aula (210), Antwerpen - Utrecht 1965

Van den Baar A.H., *Nederlands - Russisch Woordenboek*, Antwerpen - Utrecht 1994

Bańko Mirosław, *Inny słownik języka polskiego*, Wydawnictwo Naukowe PWN, Warszawa 2000

Doroszewski Witold, *Słownik języka polskiego*, Polska Akademia Nauk, Państwowe Wydawnictwo „Wiedza Powszechna”, Warszawa 1958 - 1969

Dunaj Bogusław, *Słownik współczesnego języka polskiego*, Wilga, Kraków 1996

Fisiak Jan, *Słownik polsko - angielski*, Collins - BGW, Warszawa 1996

Kopeckij L.V., Filipec J., Leska O., *Česko-ruskij slovar'*, Cechoslovackaja akademija nauk - Institut inostrannych jazykov, Moskva 1973

Martens Nico, Morciniec Elke, *Klein Nederlands - Pools en Pools - Nederlands woordenboek*, Wiedza Powszechna, Warschau 1977

Skorupko Stanisław, *Mały słownik języka polskiego*, Państwowe Wydawnictwo Naukowe, Warszawa 1969

Stanisławski Jan, *Wielki słownik polsko - angielski*, Państwowe Wydawnictwo „Wiedza Powszechna”, Warszawa 1983

Szymczak Mieczysław, *Słownik języka polskiego*, Państwowe Wydawnictwo Naukowe, Warszawa 1978 - 1981