

# Semantic Classification of Adjectives on the Basis of their Syntactical Features in Polish and English

Krzysztof Jassem (jassem@math.amu.edu.pl)

**Abstract.** The paper proposes a classification of Polish adjectives on the basis of their syntactical features in Polish and English. The classes are characterised semantically. The aim of the classification – in future research also for other parts of speech – is to assure the completeness and the homogeneity of descriptions of the entries in bilingual dictionaries appropriated for Machine Translation. The classification should prove helpful in assigning descriptions to adjectives in larger dictionaries than that discussed in the paper.

**Keywords:** bilingual lexicons, rule-based Machine Translation, adjectives, collocations

## 1. Motivation for classification

A rule-based computer analysis of a natural language assumes the existence of a lexicon in which content words are characterized by their admissible collocations with other parts of the sentence. Some lexicons include syntactical relations only, others aim at pointing out some semantic features of the related phrases. Here, we shall use the term **syntactic-semantic** collocations to denote possible coexistence in a sentence of a given word with expressions of a definite grammatical category and a definite semantic value. For example, the Polish verb *dać* (*to give*) forms (in its basic meaning) syntactic-semantic collocations with a nominal non-abstract phrase in the accusative and a nominal human phrase in the dative.

In most NLP applications, descriptions of lexical entries include collocations with subordinate categories only (a counter-example may be a so-called Link Grammar, see (LG, 1997)). That approach is followed here. The set of admissible subordinate syntactic-semantic collocations for a given word will be called here **syntactic-semantic features** of the word.

Only basic semantic features are considered here: the objects are separated into four exclusive classes: human, animate-non-human, concrete, and abstract (in fact, also a subclass of abstract terms, called 'measure', is referred to while characterising unmarked spatial adjectives).

The aim of a longer-term study is to systematise Polish content words according to their syntactic-semantic features. The hitherto existing

approaches (see (Polański, 1992), (Jassem, 1998)) have mostly used the method of defining syntactic-semantic features of each separate lexeme. Here, it is postulated to formulate – for each category of content words (i.e. verbs, nouns, adjectives and adverbs) – a classification based on syntactic-semantic features. Each class should include units of identical types of collocations.

A unit of the postulated classifications would be a lexeme and its meaning. For example, the Polish verb *uwagać* has five basic meanings: *to pay attention to something*, *to be careful*, *to beware of something*, *to think (that)*, *to consider someone as someone*. Each of the meanings forms different collocations and therefore should be assigned to a different class.

The meaning of the lexeme is proposed here to be represented by an appropriate English equivalent. Moreover, one class will include only such lexemes as have identical collocations both in Polish and English.

The direct reason for such an approach is the aim of applying the ideas to automatic translation from Polish into English. The classifications allow for an alternative of collocations on the Polish part but requires a unique collocation on the English part. This univocal character of target-language collocations aims at assuring the determinism to the translation process based on the classifications.

It is believed, however, that the classifications may prove helpful in other applications that involve the processing of Polish or English.

The benefits of existence of such classifications may be formulated as follows:

- Assuring the homogeneity of description for units of the same type. Hitherto, units forming the same collocations have quite commonly been described in a different manner because of the limitations of an analysed corpus or individual preferences of lexicographers.
- Assuring the completeness of the description.
- Facilitating the lexicographic work – in future, instead of studying a corpus a lexicographer will be able to place an entry into an appropriate class (e.g. on the basis of analogy to entries already described).
- Opening the possibility for designing algorithms based on classes of words rather than single words.

It is believed that the condition of the identity of collocations in two languages not belonging to the same family (Slavonic, German) makes each class tend to include “similar” words. For example, two

Polish verbs *postawić* (*put*) and *zamienić* (*change*) have (in their most popular usage) identical collocations in Polish (concrete noun in the accusative, PP(preposition *na* + concrete noun in the locative)) although they denote dissimilar activities. They are however, distinguished by their second collocation-argument in English, i.e. PP(*on*) and PP(*into*) respectively.

The following hypothesis is advanced: Individual classes created in the way described above lend themselves to some semantic characterisation. In this paper, the hypothesis will be verified for adjectives.

It is worth noting that a classification of adjectives according to their syntactic collocations in Polish has been proposed in (Szupryczyńska, 1980). That attempt proved very helpful for formulating the ideas presented here.

The set of adjectives under consideration is extracted from the dictionary of the POLENG system<sup>1</sup>. The set was enriched by some peculiar adjectives noted in (Szupryczyńska, 1980). The Polish collocations were verified against the corpus of Polish texts collected especially for the purpose of creating the POLENG dictionary (see (Graliński, 2000)). The English collocations were verified in (Benson et al., 1998) and BNC.

## 2. Collocations of adjectives

The following grammatical categories may be regarded as subordinate to adjectives and thus be taken into consideration as collocations in the present investigation:

- nominal phrases in a given case,
- prepositional phrases with a given preposition and a given case of a nominal part,
- some types of subordinate clauses.

A detailed list of collocations is given in section 10.

The collocations characteristic for comparative and superlative degrees of adjectives are not discussed here.

## 3. The primary division of adjectives

In (Fellbaum et al., 1993) the adjectives have been divided into four exclusive classes:

- relational adjectives that “relate/pertain to or are associated with some noun (e.g. *atomic* – *atom*, *presidential* – *president*, *dental* – *tooth*, *fraternal* – *brother*)”,
- colour adjectives,
- reference-modifying: “some refer to temporal status of the noun (former, present), others have an epistemological flavour (potential, reputed), others are intensifying (mere, sheer, actual)”,
- descriptive: “ascribe a value of an attribute to a noun (e.g. *low* for the attribute HEIGHT, *heavy* for the attribute HEIGHT)”.

The first two groups of adjectives are not investigated here. It is assumed that they do not form collocations with subordinate phrases. The reference-modifying adjectives are not excluded from the present description. Some of them have the feature of ascribing definiteness to the nominal groups they refer to (e.g. *były*: *former*). Those adjectives are treated as a separate class in the description.

#### 4. Antonymy

Mostly, antonym pairs of adjectives belong to the same class. For example, the class characterised as “readiness to execute an activity” includes both the adjectives: *chętny*: *willing*, and *niechętny*: *reluctant*. This rule has exceptions (e.g. adjectives: *zależny*: *dependent* (class 37), *niezależny*: *independent* (class 38)), and therefore for each adjective, if its antonym is included in the studied lexicon, then it is listed separately.

#### 5. Markedness

In (Fellbaum et al., 1993) the problem of markedness of adjectives is discussed. Here, only an example from that study is mentioned: The adjective *long* is unmarked with any specific semantic value and the sentence *The road is X miles long* is acceptable. The adjective *short* is marked with some semantic interpretation (length of the road should be less than, say, 1 mile) and the sentence *The road is X miles short* is unacceptable. Unmarked spatial adjectives form characteristic collocations both in Polish in English. They are listed in group 5 of the present classification.

Table I. Syntactic-semantic disambiguation of adjectives

| adjective+sense             | class | adjective+sense         | class |
|-----------------------------|-------|-------------------------|-------|
| <i>ciekawy: interesting</i> | 18.2  | <i>ciekawy: curious</i> | 46    |
| <i>jasny: clear</i>         | 17    | <i>jasny: bright</i>    | 56    |
| <i>solidny: reliable</i>    | 2.4   | <i>solidny: solid</i>   | 56    |
| <i>surowy: severe</i>       | 1.1   | <i>surowy: raw</i>      | 56    |
| <i>wolny: slow</i>          | 2.4   | <i>wolny: free</i>      | 38    |

## 6. Polysemy

In (Fellbaum et al., 1993) it is mentioned that polysemy of adjectives may be solved by means of finding different antonyms to different senses of an adjective (e.g. *fresh – stale*, *fresh – preserved*).

Here, it is suggested that the syntactic-semantic classification may also help disambiguate senses of polysemous adjectives. The table I shows some examples.

## 7. Predicative and attributive use

When speaking of syntactic characteristics of adjectives, the problem of attributive versus predicative use is commonly treated as primary. Here, the problem is disregarded (this does not mean that it is not important for Polish-English MT in general). It is assumed that the collocations mentioned in the classification refer to both uses (although in real texts the collocations mostly occur when adjectives are used predicatively, it is still possible to generate, for each type of a collocation, an acceptable sentence in which the collocation coexists with an attributive use of the adjective).

## 8. Facultativity of collocations

In (Szupryczyńska, 1980), the problem of facultativity of categories subordinate to adjectives is discussed. The author claims that in most cases the occurrence of subordinate categories is not obligatory. This idea is, on principle, followed here. When an adjective is classified to a class with non-zero collocations (i.e. any class other than class 56), then it is assumed that it may appear in an acceptable sentence with or without the collocations.

This general rule has an exception. When the same Polish adjective has a few senses and one of them belongs to class 56, then the collocations of other senses are regarded as obligatory. For example, the adjective-sense *głodny: hungry* belongs to class 56 and the adjective-sense *głodny: avid* belongs to class 47, which means that it connotes an abstract nominal phrase in the genitive. The occurrence of such a phrase in an expression is obligatory for the Polish adjective *głodny* to be interpreted as *avid* rather than *hungry*.

## 9. The boundaries of connotations

In any attempt to form the relations between a governing and a governed component, a natural question arises: Where is the boundary between treating a component as subordinate to other component and regarding it as free (not governed). This problem concerns mainly prepositional phrases and subordinate clauses (rather than nominal phrases). Usually, this boundary is defined on the basis of linguistic intuition of researchers. The bi-language approach may enhance the human intuition in the following way: “if in doubt, regard the component as free, if its transfer from one language to another is independent of what appears to be the governing word”.

## 10. Classification of adjectives

For the sake of better readability, the classes of adjectives have been ordered in 8 groups (each group is assigned some rather vague semantic interpretation). The next (and crucial) stage of classification divides adjectives into 56 classes. A few classes consist of adjectives of similar meaning which form collocations differing slightly in one of the two languages. Those classes are divided into sub-classes. Altogether 68 sub-classes are distinguished. Each sub-class forms syntactic-semantic collocations specific to itself.

For each sub-class, several examples of adjective units, extracted from the dictionary of POLENG system, are given. The examples are generally listed in the alphabetical order with the exceptions for the antonym pairs which are listed directly one after another.

Some adjective units do not lend themselves to the semantic characterisation: they belong to a certain (sub)class as far as syntax is concerned but they do not share the semantic characteristics. Such adjectives are mentioned at the end of the list for the (sub)class and are marked with italic font.

For each (sub)class, admissible collocations of adjectives, both in Polish, in the left column, and English, in the right column, are listed. The collocations are encoded with the following formalism:

- lower case words denote prepositions,
- upper case words denote categories:
  - N, G, D, A, I, L stand for nominal phrases in an appropriate case (for Polish),
  - NP = any nominal phrase (for English),
  - H (human), An (animate), Ab (abstract), Con (concrete) denote semantic constraints on nominal groups, e.g. G:Ab denotes an abstract nominal phrase in the genitive. Semantic constraints may be preceded by the character '-', which stands for complementation of a semantic class (e.g. -H = non-human)
  - GR = gerund form. For the Polish part the symbol is preceded by the case that the gerund form should occur in,
  - INF = bare infinitive, to+INF = *to* + infinitive,
  - ŻE, CZY, BY denote specific subordinate clauses in Polish, i.e. clauses starting with the conjunctions ŻE (English standard equivalent: *that*), CZY (English equivalent: *if*) or BY (English equivalent: *in order to*) respectively,
  - *za to* BY etc. denote constructions specific of Polish, in which a nominal phrase is substituted by a subordinate clause,
- brackets denote alternative constructions for Polish, which have the same equivalents in English,
- digits separate alternative constructions for both languages,
- the colon (;) separates descriptions of more than one collocation (in rare situations when an adjective connotes more than one construction).

## Group I. Adjectives usually referring to humans

### Class 1. Attitude towards other people

#### Subclass 1.1. High expectations

dla [wobec, względem, w stosunku do] G:H          with H

bezwzględny: strict, groźny: severe, ostry: severe, rygorystyczny: strict, srogi: severe, surowy: severe, stanowczy: firm, szorstki: curt

*bezczelny: cheeky, cierpliwy: patient, niecierpliwy: impatient, uczciwy:*

*honest, nieuczciwy: dishonest*

**Subclass 1.2.** Anger

- |                                |                    |
|--------------------------------|--------------------|
| 1) na A:H ; za A:Ab [o A: Ab]  | 1) with H ; for Ab |
| 2) na A:H ; za to ŻE [o to ŻE] | 2) with H ; for GR |

rozzłoszczony: irritated, wściekły: furious, zagniewany: angry, zły: angry

The following example should clarify the notation of the second type of the above collocations – Polish: *On był wściekły na mnie, za to że się spóźniłem*, English: *He was angry with me for being late*.

**Subclass 1.3.** Cool negative attitude

dla [wobec, względem, w stosunku do] G:H of H

krytyczny: critical, sceptyczny: sceptical, wymagający: demanding

The two remaining subclasses are described as 'other adjectives describing attitude'. Subclass 1.5 is distinguished from subclass 1.4 by the ability of Polish adjectives to connote nominal groups in dative.

**Subclass 1.4.** Other adjectives describing attitude

dla [wobec, względem, w stosunku do] G:H to H

agresywny: aggressive, arogancki: arrogant, bezlitosny: merciless, litościwy: merciful, brutalny: brutal, czuły: affectionate, nieczuły: heartless, dobry: nice, niedobry: nasty, fajny: great, groźny: threatening, niegroźny: unalarming, grubiański: rude, grzeczny: polite, niegrzeczny: impolite, hardy: arrogant, hojny: generous, skąpy: stingy, kochany: dear, koleżeński: amicable, niekoleżeński: uncomradely, krzywdzący: unjust, sprawiedliwy: just, łagodny: gentle, łaskawy: gracious, miłosierny: merciful, miły: kind, niemiły: unkind, niebezpieczny: dangerous, nieufny: distrustful, obojętny: indifferent, obraźliwy: offensive, okrutny: cruel, pobłażliwy: indulgent, przyjacielski: friendly, przyjemny: pleasant, nieprzyjemny: unpleasant, serdeczny: cordial, sprawiedliwy: fair, niesprawiedliwy: unfair, sympatyczny: nice, niesympatyczny: impolite, szczery: sincere, nieszczery: unsincere, uczynny: obliging, ugrzeczniony: attentive, usłużny: submissive, wredny: vicious, wrogi: hostile, złośliwy: malicious

**Subclass 1.5.**

dla [wobec, względem, w stosunku do] G:H [D] to H

lojalny: loyal, nielojalny: disloyal, pomocny: helpful, przychylny: friendly, nieprzychylny: unfriendly, przyjazny: friendly, nieprzyjazny: unfriendly, przykry: unpleasant, wierny: faithful, niewierny: unfaithful, życzliwy: kindly, nieżyczliwy: unkindly

**Class 2.** A human-activity relationship**Subclass 2.1.** Readiness (expression of will) to execute an activity

- |  |           |
|--|-----------|
| 1) na A:Ab [do G:Ab]                         | 1) for Ab |
| 2) na A:GR [na to BY] [do G:GR] [do tego BY] | 2) to+INF |
| 3) INF [BY]                                  | 3) to+INF |



Most of the below adjectives collocate with both PP constructions mentioned in 1) and 2) (i.e. those starting with prepositions *na* or *do* respectively). Yet, some of them show clear preference to co-occur with only first or second type of phrases. Those are marked with (*na*) or (*do*) denoting the only prepositions that the adjectives are likely to coexist with.

chętny: willing, niechętny: reluctant, gotowy: ready, niegotowy: unready,  
ochoczy: eager, skory: willing, skłonny(do): prone, zdecydowany(na):  
strong-minded, zdesperowany(na): desperate, zdeterminowany: determined

**Subclass 2.2.** Qualification to execute an activity

- |                         |           |
|-------------------------|-----------|
| 1) do G:Ab              | 1) for Ab |
| 2) do G:GR [do tego BY] | 2) to+INF |
| 3) INF [BY]             | 3) to+INF |

kompetentny: competent, niekompetentny: incompetent, odpowiedni:  
adequate, nieodpowiedni: inadequate powołany: competent, niepowołany:  
not competent, upoważniony: authorised, nieupoważniony: unauthorized,  
uprawniony: entitled, nieuprawniony: unauthorized, uprzywilejowany:  
privileged, wykwalifikowany: qualified, zobowiązany: obliged

**Subclass 2.3.** Quality of execution

w L:Ab [z G:Ab] at Ab

biegły: proficient, dobry: good, zły: poor, doświadczony: experienced,  
nie doświadczony: inexperienced, genialny: brilliant, kiepski: poor, marny:  
poor, niezły: pretty good, pomysłowy: ingenious, sprytny: clever, świetny:  
excellent, wprawny: skilful, wszechstronny: versatile, wybitny: outstanding,  
wydajny: efficient, zgrabny: deft, zmyślny: clever, zręczny: skilful

**Subclass 2.4.** Attitude while executing an activity, efficiency of execution

w L:Ab in Ab

aktywny: active, bierny: passive, bez troski: careless, bezwzględny: ruthless,  
biegły: fluent, cierpliwy: enduring, niecierpliwy: impatient, czynny: active,  
dociekliwy: inquiring, dogłębny: profound, dokładny: accurate, niedokładny:  
inaccurate, dynamiczny: dynamic, dyskretny: discreet, konsekwentny:  
consistent, niekonsekwentny: inconsistent, liberalny: liberal, mocny: strong,  
słaby: weak, niechlujny: sloppy, niezastąpiony: irreplaceable, nie równany:  
matchless, precyzyjny: accurate, nieprecyzyjny: inaccurate profesjonalny:  
professional, nieprofesjonalny: unprofessional, przodujący: leading, regularny:  
regular, nieregularny: irregular, rzetelny: reliable, nierzetelny: unreliable,  
silny: strong, skuteczny: effective, bezskuteczny: ineffective, nieskuteczny:  
ineffective, solidny: reliable, niesolidny: unreliable, umiarkowany: moderate,  
uporczywy: persistent, wiarygodny: credible, wpływowy: influential,  
wprawny: practised, wytrwały: persistent, zaawansowany: advanced

**Subclass 2.5.** Careful approach to activity, persistence, common sense

przy L:Ab [w L:Ab] about Ab

niedbały: casual, opieszwały: negligent, ostrożny: careful, praktyczny: practical, precyzyjny: precise, realistyczny: realistic, rozsądny: reasonable, rozważny: cautious, nierozważny: reckless, skrupulatny: precise, subtelny: subtle, staranny: careful, uparty: stubborn, uważny: careful, wybredny: choosy, zacięty: obstinate,  
*skromny: modest*

**Class 3.** Resilience, sensitivity

- |             |          |
|-------------|----------|
| 1) na A:Ab  | 1) to Ab |
| 2) na to ŻE | 2) to GR |

czuły: sensitive, nieczuły: insensitive, obojętny: insensitive, odporny: resistant, odporny: immune, nieodporny: susceptible, oporny: resistant, podatny: susceptible, uodporniony: immune, wrażliwy: sensitive, niewrażliwy: insensitive, wytrzymały: resilient  
*głuchy: deaf, ślepy: blind*

**Class 4.** Satisfaction

- |             |            |
|-------------|------------|
| 1) I, [z G] | 1) with Ab |
| 2) ŻE       | 2) THAT    |

sfrustrowany: frustrated, usatysfakcjonowany: satisfied, nieusatysfakcjonowany: unsatisfied, zadowolony: pleased, niezadowolony: unsatisfied

**Class 5.** Happiness

- |    |      |
|----|------|
| ŻE | THAT |
|----|------|

smutny: sad, szczęśliwy: happy,

**Class 6.** Confidence, awareness

- |       |          |
|-------|----------|
| 1) G  | 1) of NP |
| 2) ŻE | 2) THAT  |

pewny: sure, niepewny: uncertain, niepewny: doubtful, świadomy: aware, nieświadomy: unaware

**Class 7.** A state being a consequence of an event

**Subclass 7.1.** Responsibility

- |      |        |
|------|--------|
| za A | for NP |
|------|--------|

odpowiedzialny: responsible, skazany: convicted

**Subclass 7.2.** Gratefulness

- |          |             |
|----------|-------------|
| D:H za A | to H for NP |
|----------|-------------|

wdzięczny: grateful, zobowiązany: thankful

**Class 8.** Fame

- |              |        |
|--------------|--------|
| 1) z G:Ab    | for Ab |
| 2) z tego ŻE | for GR |

sławny: famous, słynny: famous, znany: well-known

**Class 9. Abundance**

w A in NP  
 bogaty: rich, obfity: rich, ubogi: poor, zasobny: wealthy

**Class 10. Talents**

do G:Ab at Ab  
 uzdolniony: talented: utalentowany: gifted

**Class 11. Attitude while judging**

do [wobec, względem, w stosunku do] G about NP  
 obiektywny: objective, subiektywny: subjective, sceptyczny: sceptical

**Class 12. Deserving**

1) G of NP  
 2) INF [G:GR] GR  
 3) tego BY GR

godny: worth, niegodny: unworth, wart: worth, niewart: unworth

**Class 13. Mindfulness**

1) o A 1) about NP  
 2) o to BY [o to ŻE] 2) about GR

dbały: solicitous, niespokojny: uneasy, spokojny: unworried, troskliwy:  
 solicitous, zazdrosny: jealous

**Class 14. Guilt**

1) G:Ab [D:Ab] of Ab  
 2) tego ŻE of GR

winny: guilty, niewinny: innocent

**Class 15. Attack – defence**

wobec [w stosunku do] G against NP

bezbronny: defenceless, gwałtowny: violent

**Group II. Adjectives describing subjective characteristics of non-human objects**

This group includes adjectives, which usually refer to non-human objects. The collocations refer to either human estimation of the characteristic (classes 17, 18.1, 18.2), or human activity (class 16).

**Class 16. Usefulness for an activity**

Some of the adjectives in this class tend to form collocations in Polish with PPs(do) (*pożyteczny: useful*), some accept also PPs(na) (*przydatny: useful*). Here, the two types are not separated.

1) na A:Ab [do G:Ab] 1) for Ab  
 2) w L:Ab 2) in Ab

- |   |                   |
|---|-------------------|
| 3) na A:GR [na to BY] [do G:GR] [do tego BY]                    | 3) to+INF         |
| 4) INF [BY]   | 4) to+INF         |
| 5) dla G:H [D:H] ;<br>na A:GR [na to BY] [do G:GR] [do tego BY] | 5) for H ; to+INF |
| 6) dla G:H [D:H] ; BY   | 6) for H ; to+INF |

adekwatny: adequate, dogodny: convenient, doskonały: perfect, dostateczny: sufficient, niedostateczny: insufficient, fundamentalny: fundamental, idealny: perfect, istotny: essential, konieczny: necessary, kluczowy: fundamental, możliwy: possible, niemożliwy: impossible, newralgiczny: critical, nieodzowny: indispensable, niewystarczający: insufficient, niezastąpiony: indispensable, niezbędny: necessary, odpowiedni: fit, nieodpowiedni: unsuitable, podstawowy: essential, potrzebny: necessary, niepotrzebny: unnecessary, pożądany: desirable, niepożądany: undesirable, pożyteczny: useful, przydatny: useful, nieprzydatny: useless, stosowny: appropriate, niestosowny: inappropriate, użyteczny: useful, bezużyteczny: useless, właściwy: suitable, niewłaściwy: unsuitable, wygodny: convenient, niewygodny: inconvenient, zasadniczy: fundamental, zbędny: unnecessary, niezbędny: necessary, zdatny: suitable, niezdatny: unfit

Classes 17 and 18 both include adjectives, which describe a characteristic of an (usually non-human) object relative to a certain (usually human) subject. The relation between “an object” and “a subject” is described in Polish by the preposition *dla*. In English the relation is coded by either the preposition *to* or *for*. The distinction between the two classes lends itself subjects to any semantic description. The below attempt of such a semantic distinction may be treated only as a weak guidance.

### Class 17. Perception, comprehensibility

|         |      |
|---------|------|
| dla G:H | to H |
|---------|------|

banalny: trivial, bezproblemowy: troubleless, dostępny: available, niedostępny: unavailable, dostrzegalny: perceptible, niedostrzegalny: imperceptible, ewidentny: evident, jasny: clear, niejasny: unclear, klarowny: clear, przejrzysty: clear, oczywisty: obvious, odczuwalny: noticeable, nieodczuwalny: unnoticeable, osiągalny: attainable, nieosiągalny: unattainable, słyszalny: audible, niesłyszalny: inaudible, tajemniczy: obscure, widoczny: visible, niewidoczny: invisible, wyobrażalny: imaginable, niewyobrażalny: unimaginable, zauważalny: perceptible, niezauważalny: imperceptible, zrozumiały: comprehensible, niezrozumiały: incomprehensible  
*akceptowalny: acceptable, atrakcyjny: attractive, bezcenny: priceless*

### Class 18. Other “subjective” adjectives

Subclass 18.1 is distinguished from subclass 18.2 by the ability of connotating two arguments; the second argument describes an activity.

#### Subclass 18.1. Difficulty of activity

dla G:H ; do G:GR for H ; to+INF

banalny: trivial, łatwy: easy, prosty: simple, skomplikowany: complex,  
trudny: difficult

### Subclass 18.2.

dla G:H for H

bezwartościowy: worthless, błaży: trivial, bolesny: painful, cenny: valuable,  
ciekawy: interesting, nieciekawy: uninteresting, czytelny: readable,  
nieczytelny: illegible, dochodowy: profitable, niedochodowy: unprofitable,  
dokuczliwy: troublesome, doniosły: important, dotkliwy: painful, dozwolony:  
admissible, niedozwolony: prohibited, fatalny: fatal, idealny: ideal,  
interesujący: interesting, nudny: boring, istotny: significant, nieistotny:  
insignificant, kłopotliwy: troublesome, korzystny: beneficial, korzystny:  
favourable, niekorzystny: unprofitable, marginalny: marginal, nieoceniony:  
invaluable, nieoczekiwany: unexpected, obowiązkowy: compulsory, opłacalny:  
profitable, nieopłacalny: unprofitable, owocny: profitable, pamiętny:  
memorable, pociągający: attractive, prestiżowy: prestigious, przełomowy:  
momentous, przewidywalny: predictable, nieprzewidywalny: unpredictable,  
przystępny: intelligible, rentowny: profitable, uciążliwy: arduous, udany:  
successful, wartościowy: valuable, ważny: important, nieważny: unimportant,  
wygodny: comfortable, niewygodny: uncomfortable, wykonalny: feasible,  
niewykonalny: unfeasible, znaczący: significant

### Group III. Adjectives describing relations between two non-abstracts objects

#### Class 19. Separation

od G:-Ab from -Ab

oddzielny: separate, odosobniony: isolated, odrębny: separate, odseparowany:  
separate

#### Class 20. Distance

od G:-Ab o Dist from -Ab by Dist

daleki: far, oddalony: distant, odległy: distant

#### Class 21. Geometrical relationship

do [wobec, względem, w stosunku do] G to NP

prostopadły: perpendicular, przeciwległy: opposite, przeciwny: opposite,  
przyległy: adjacent, sąsiadujący: adjacent, styczny: adjacent, równoległy:  
parallel, zbieżny: convergent

### Group IV. Adjectives referring to various types of objects

For adjectives in group IV, it is hard to characterise the semantics of objects they usually refer to.

**Class 22.** Hierarchy, obedience

D [dla G, wobec G, względem G, w stosunku do G] to NP

nadrzędny: superior, podległy: subordinate, podrzędny: subordinate,  
 podwładny: subordinate, compliant, równorzędny: equivalent, posłuszny:  
 obedient, nieposłuszny: disobedient

**Class 23.** Familiarity

Subclass 23.1 includes adjectives which, in Polish, may collocate with either PPs(dla) or NPs(Dative). For adjectives in subclass 23.2, only NPs(Dative) are admissible.

**Subclass 23.1.**

dla G [D] to NP

bliski: close, obcy: alien, obojętny: unimportant, wiadomy: known,  
 niewiadomy: unknown, znajomy: familiar, nieznaomy: unknown, znany:  
 known, nieznan: unknown  
*specyficzny: specific*

**Subclass 23.2.**

D to NP

drogi: dear, miły: dear, pokrewny: related  
*należny: due, pozostały: left, przeciwny: opposed*

**Class 24.** Comparison

Some adjectives describing comparison may in Polish optionally connotate a nominal phrase in dative. They are marked with (D).

**Subclass 24.1.** Relation of similarity

z I with NP

bliskoźnacny(D): synonymous, dobrany: well-matched, jednoczesny:  
 simultaneous, kompatybilny: compatible, niekompatybilny: incompatible,  
 porównywalny: comparable, równoczesny: concurrent, spójny: compatible,  
 niespójny: incompatible, sprzeczny: inconsistent, synonimiczny: synonymous,  
 współbieżny: concurrent, współczesny(D): contemporary, zbieżny:  
 convergent, zespolony: joint, zgodny: compatible, zgodny(D): consistent,  
 niezgodny: inconsistent  
*obyty: familiar, zaprzyjaźniony: friendly, zaznajomiony: familiar, rozbieżny:  
 divergent*

**Subclass 24.2.** Relation of identity

z I to NP

identyczny: identical, jednakowy: identical, jednoznaczny(D): equivalent,  
 niezgodny: contrary, równoprawny(D): equal, równy(D): equal, nierówny:  
 unequal, równoznaczny(D): tantamount, sprzeczny: contradictory,  
 tożsamy(D): identical  
*wspólny(D): common*

Additionally, the adjective identyczny connotates PPs(jak+Nom).

**Subclass 24.3.** Other relations of comparison

do [wobec, względem, w stosunku do] G to NP  
sporadically z I

alternatywny: alternative, analogiczny: analogous, odwrotny: contrary,  
odwrotny: opposite, podobny: similar, niepodobny: dissimilar,  
porównywalny: comparable, nieporównywalny: incomparable,  
proporcjonalny: proportional, przeciwny(D): contrary, równoważny(D):  
equivalent, zbieżny: convergent, zbliżony: similar

**Class 25.** Typicality of an object relative to a class it belongs to

dla G of NP

charakterystyczny: characteristic, reprezentatywny: representative, typowy:  
typical, nietypowy: atypical, znamieny: characteristic

**Class 26.** Harmfulness

dla G to NP

groźny: dangerous, niegroźny: harmless, szkodliwy: harmful, nieszkodliwy:  
harmless, śmiertelny: deathly, zabójczy: deadly

**Class 27.** Difference

od G [niż N] from NP

inny: different, odmienny: different, różny: different

**Class 28.** Popularity

wśród [pośród] G:H with H

popularny: popular, głośny: popular, modny: popular

**Class 29.** Safety

przed L:Ab from Ab

bezpieczny: safe, zabezpieczony: secure

**Class 30.** Quantification

z G [spośród G] of NP

jeden: one, każdy, each, żaden: none, wszyscy: all

**Group V.** Unmarked spatial adjectives**Class 31.**

The class contains adjectives which in Polish occur in sentences of the form *jest(is) ADJ na Measure* (e.g. *ciasto jest grube na dwa palce – the pastry is two fingers thick*).

na Measure Measure

długi: long, gruby: thick, szeroki: wide, rozległy: wide, wysoki: high, wysoki:  
tall

**Class 32.**

The class contains adjectives which in Polish occur in sentences of the form *jest(VERB is) ADJ o(PREP) Measure*.

o Measure

Measure

daleki: distant, odległy: distant, opóźniony: late, opóźniony: backward, spóźniony: late

**Group VI. Deverbal adjectives (participles)**

In lexicography, adjectival participles are treated inconsistently – sometimes they form separate entries and sometimes they function only as inflected forms of verbs. The discussion on the criteria of categorisation of adjectival participles in Polish may be found in (Węgrzynek, 1995). The question arises whether syntactic-semantic features of adjectival participles should be characterised on the background of adjectives or verbs. The attributive role of participles would call for the first approach. The “inheritance” of collocations from verbs that participles are derived from would suggest the other approach.

The criterion for categorisation of participles, assumed here, is strongly connected with Polish-English translation. The participle is treated as an inflected form of the verb (and therefore is excluded from the current investigation concerning adjectives) if the rule-based, compositional translation of both the participle and its complements, on the basis of lexical information attached to the verb, would give a correct result.

Thus, a participle submits to the hitherto classification if one of the two following cases takes place:

1. The English equivalent of the Polish participle is not a participle form of the corresponding English verb (e.g. *zależny: dependent*). Such participles are treated in this classification as regular adjectives.
2. The collocations of the participle (in either Polish or English) differ from that of the verb. This takes place for participles listed below. For example, (Benson et al., 1998) does not list the PP *at NP* as a collocation of the verb *surprise* but the same PP is mentioned as a collocation of the participle *surprised*.

Classes 33 – 36 deal with the collocations of adjectival participles, which are not inherited from the verbs.

**Class 33.** Participles expressing strong emotions

I: Ab

at Ab

przejęty: excited, przerażony: terrified, zdumiony: amazed



**Class 34.** Participles expressing boredom

I: Ab of Ab  
 zmęczony: tired, znudzony: bored, znużony: weary

**Class 35.** Participles expressing care

I: Ab about Ab  
 przejęty: worried, zatroskany: worried

**Class 36.** Other participles

I by NP  
 ograniczony: limited

**Group VII. Other classes of collocations**

Group VII includes adjectives whose syntactic-semantic features form small (usually one-element) classes. The classes are not characterised semantically here. Perhaps an analysis of a larger dictionary would allow for the extension of the classes with other adjectives and consequently lead to their semantic interpretation.

**Class 37.**

1) od G 1) on NP  
 2) od tego ŻE 2) on GR  
 zależny: dependent

**Class 38.**

od G of NP  
 niezależny: independent

**Class 39.**

od G [z G] from NP  
 pochodny: derived, powstały: originating

**Class 40.**

do G:Ab on Ab  
 zapalony: keen

**Class 41.**

z I:-H from -H  
 rozdzielny: separate

**Class 42.**

na L:Ab at Ab  
 obecny: present

**Class 43.**

na L:Ab from Ab  
nieobecny: absent

**Class 44.**

na A:Ab against Ab  
skuteczny: efficacious

**Class 45.**

1) do G:Ab 1) of Ab  
2) IN [BY+IN] [do G:GR] 2) of GR  
zdolny: capable

**Class 46.**

1) G:Ab 1) about Ab  
2) tego CZY 2) about GR  
ciekawy: curious

**Class 47.**

G:Ab of Ab  
pełen: full, głodny: avid

**Class 48.**

G:Con with Con  
pełen: filled, syty: well-fed

**Class 49.**

z I: Con of Con  
rozłączny: exclusive

**Class 50.**

1) z I 1) of NP  
2) ŻE 2) THAT  
3) z G:GR 3) TO+INF  
dumny: proud

**Class 51.**

w L:H with H  
zakochany: in love

**Class 52.**

o L:Ab about Ab  
o tym ŻE about GR  
przekonany: convinced, przeświadczony: convinced, rozstrzygający: decisive

**Class 53.**

z G: Ab with Ab  
 słaby: weak

**Class 54.**

po L after NP  
 pozostały: left

**Group VIII. Adjectives which do not connote subordinate phrases**

This most numerous group of adjectives does not submit to any semantic interpretation because of wide variation of adjectives it includes. The syntactical interpretation is following: the adjectives in the positive degree do not connote any definite grammatical categories.

As mentioned in section 3, amongst those adjectives, a class of reference-modifying adjectives, which assign definiteness to nouns they refer to, is selected.

**Class 55.** Reference-modifying “definite” adjectives

były: former, kolejny: next, miniony: past, następny: following, przyszły: next, następujący: successive, niniejszy: present, obecny: present, niższy: following, poprzedni: preceding, powyższy: above-mentioned, przeszły: past, przyszły: next, rzeczone: said, ubiegły: past, uprzedni: prior, wczorajszy: yesterday, zeszyły: last,  
*całkowity: entire, cały: whole*

**Class 56.** Other adjectives

Below, a few first (in the alphabetical order) adjectives of the POLENG dictionary, which belong to class 56, are listed:

abonamentowy: subscription, abonencki: subscriber, absolutny: absolute, abstrakcyjny: abstract, absurdalny: absurd, achromatyczny: achromatic, adaptacyjny: adaptive, addytywny: additive, administracyjny: administrative, adresowy: address, aerodynamiczny: aerodynamic, afrykański: African, akademicki: academic

**11. A classification algorithm for adjectives**

The classification presented above should facilitate the task of describing syntactic-semantic features of adjective units from beyond the lexicon cited here. The following algorithm may guide a lexicographer (possibly with the help of a computer program):

0. For a given Polish adjective, divide the adjective into units of the type: Polish adjective – English adjective according to its meanings.  
 For each adjective-unit:

1. If the adjective is relational or denotes colour, insert it to class 56. Else:
2. If the adjective is reference-modifying, insert it to class 55 – if it assign definiteness to the nominal group – or to the class 56 otherwise. Else:
3. If the adjective is a participle, check if it fits to classes 33-36. Else:
4. If the adjective is unmarked spatial, check if it fits to classes 31-32. Else:
5. Try to classify the adjective to one of the classes 1-30 by means of semantic analogy. Else:
6. Try to classify the adjective to one of the classes 37-54 by means of syntactic analogy. Else:
7. If the adjective has subordinate collocations, describe the collocations by means of a given formalism. Else:
8. Insert the adjective to class 56.

## 12. Conclusions

The paper presents a classification of adjectives according to their syntactic-semantic features in Polish and English. Designing such a classification aims at assuring the completeness of syntactic-semantic descriptions in a lexicon, as well as homogeneity of descriptions for similar adjectives. The classification should also prove helpful in assigning descriptions to adjectives in larger dictionaries.

The description of bilingual collocations of words seems indispensable for any lexicon-based Machine Translation systems. Such descriptions are necessary for:

- finding the dependencies in the syntactical structure of the source expression,
- calculating the best target equivalent of a source word on the basis of its actual collocations in an input expression,
- determining the correct syntactical structure of the target expression.

## Notes

<sup>1</sup> The dictionary consists of roughly 15 000 lexemes occurring most frequently in computer-oriented texts.

## References

- : 1997, 'Link Grammar'. <http://bobo.link.cs.cmu.edu>.
- Benson, M., E. Benson, and R. Ibsen: 1998, *The BBI Dictionary of English Word Combinations*. Poznań: Wydawnictwo Omnibus.
- Fellbaum, C., D. Gross, and K. Miller: 1993, 'Adjectives in WordNet'. <http://www.cogsci.princeton.edu/~wn/papers/>.
- Graliński, F.: 2000, 'Hasłowanie korpusu polskich tekstów informatycznych (1,2 mln słów) — raport (The Report on Lemmatisation and Disambiguation of a Corpus of Polish Texts Related to Computer Science)'. In: K. Jassem (ed.): *Speech and Language Technology*, Vol. 4. Poznań: Polish Association of Phonetics, pp. 128–140.
- Jassem, K.: 1998, 'Format of the Entry into the Electronic Polish-English Dictionary Designed for MT'. In: W. Jassem, C. Basztura, and K. Jassem (eds.): *Speech and Language Technology*, Vol. 2. Poznań: Polish Association of Phonetics, pp. 137–149.
- Polański, K. (ed.): 1980–1992, *Słownik syntaktyczno-generatywny czasowników polskich (The syntactic-generative dictionary of Polish verbs)*. Warszawa-Wrocław-Katowice-Gdańsk: Wydawnictwo im. Ossolińskich.
- Szupryczyńska, M.: 1980, *Opis składniowy polskiego przymiotnika (A syntactical description of Polish adjectives)*. Toruń: Polska Akademia Nauk.
- Węgrzynek, K.: 1995, 'Składnia przymiotnika polskiego w ujęciu generatywno-transformacyjnym (Syntax of Polish adjectives in the generative-transformational approach)'. In: *Prace Instytutu Języka Polskiego*. Kraków: Polska Akademia Nauk, Instytut Języka Polskiego.

*Address for Offprints:* Adam Mickiewicz University  
Faculty of Mathematics and Computer Science  
ul. Matejki 48-49  
60-769 Poznań

