Lexicalizing the Concept of Sentimentality

Željko Bujas
Faculty of Philosophy, Zagreb

The assumption that "words express ideas" is first tested against a rank list of general vocabulary of American English. The lexical items expressing the top concepts established (Man, Time, Space) are found to closely correspond to top items in a similar rank list of Russian. This proves both the universality of the top concepts and the validity of the original assumption. The paper then proceeds to investigate how the concept of "sentimentality" is lexicalized. First, definitions from obviously relevant entries (containing the base *sentiment-*) in a collegiate dictionary of English are mined for keywords. These lead to other entries and their definitions which, in their turn, yield new keywords. This production of lexical items expressing the concept of "sentimentality" is considerably expanded in the next step when two large thesauri of English are submitted to a similar process. Finally, the resulting lists are convincingly demonstrated to be rewarding devices in a (simulated) computer-supported search for the presence, direct and indirect, of the diffuse concept of "sentimentality" in such a vast and varied mine as the text of Encyclopaedia Britannica.

1.

1.1. It is probably safe to say that the average speaker of any language intuitively feels that "words express ideas". From this general statement to the more specific "vocabulary items express concepts" there is but a step.

1.2. The conceptual structure of vocabulary is most convincingly illustrated by analyzing frequency-ordered rank lists of the vocabularies of individual languages. So, for instance, the latest word count of American English (*Frequency Analysis of English Usage*, 1982)[1] reveals that the 100 most frequent words of

1. Nelson Francis and Henry Kučera, *Frequency Analysis of English Usage* (asstd. by Andrew W. Mackie), Houghton Mifflin, Boston, 1982. This is in fact the 'lemmatized' version

American English are made up largely of such nonlexical items as articles, prepositions, conjunctions and pronouns (best illustrated by the first ten: *the, be, of, and, a, in, he, to* /inf/, *have, to* /prep/). Among these 100, however, we find only one adjective (*new*), 16 verbs (in order of frequency: *be, have, do, say, make, go, take, come, see, get, know, give, find, use, think* and *seem*), a number of which are nonlexical in varying degrees, and 9 nouns (*man, time, year, state, day, way, people, Mister* and *world*).

1.3. Nouns can be said to express most immediately a language's inventory of abstract concepts. It can likewise be argued that the most frequent nouns express the most fundamental abstract concepts, thus indicating Man's fundamental preoccupations. These, according to the most frequent nine English nouns are: Man himself (*man, state, people, Mister*), Time (*time, year, day*) and Space (*world*). These are, of course, rather simplistic conclusions — not accounting completely for *way* (whose spatial meanings account for only 41% of its occurrences),[2] or *state* (lumped together with *man* and *people*, and with its meaning of "condition" /16% of occurrences/ neglected).

However, the universality of this observation is corroborated by a comparable word count of Russian,[3] where 100 top-frequency items contain 11 nouns ranked as follows: *god* (year), *delo* (affair, matter), *vremya* (time), *chelovek* (man), *lyudi* (people), *ruka* (hand), *zhyzn'* (life), *den'* (day), *tovarishch* (comrade, Mister), *rabota* (work) and *glaz* (eye). There is an uncanny closeness of rank positions between the Russian *tovarishch* (92nd out of 100) and its cultural equivalent in American English *Mister* (94th).

1.4. For the purpose of this paper, which is to investigate the lexicalization of the concept of "sentimentality", it is not necessary to go further down the said rank list of American English (and learn that the word *sentimentality* occurs only once in a one-million-word corpus, *sentimental* 14 times and *sentiment* 25 times).

Equally, we do not need to involve ourselves with such semantic issues as the distinction between the designatum of a concept (e.g. "cathood"), its denotatum (the actual cat seen) and its linguistic expression (*cat*).

We will, for our purpose, simply assume that, to the degree that words express concepts (which they obviously do), we can profitably investigate the phenomenon of the lexicalization of a specific concept with the aid of dictionaries viewed as inventories of lexically expressed concepts.

of Francis and Kučera's earlier frequency analysis of the well-known Brown Corpus (*Computational Analysis of Present-Day American English*, Brown Univ. Press, Providence, R.I., 1967). A British-English word count, closely modeled on Francis and Kučera's work is Knut Hofland and Stig Johansson's *Word Frequencies in British and American English* (The Norwegian Computing Centre for the Humanities, Bergen, 1982).

2. Based on the semantic breakdown in Michael West's *A General Service List of English Words*, Longmans, 1959.

3. L. N. Zasorina (ed.), *Chastotnyj slovar' russkogo jazyka* (Russkij jazyk, Moscow, 1977).

2.

2.1. Webster's Ninth Collegiate (1983) — MW9 for short — offers the following set of lexically expressed concepts, in the form of entries, centered around "sentimentality": first *sentimentality* itself, flanked by *sentiment, sentimental, sentimentalism* and *sentimentalize*, plus the run-on entries *sentimentally, sentimentalist* and *sentimentalization*. Each of these — except for the undefined run-on entries — contains a description (definition) of its meaning. This is either a simple description (like *sentimentalize*, broken down only into a *vi* and a *vt* definition), or one with numbered senses (*sentimentalism, sentimentality*), or one with both numbered senses and lettered subsenses (*sentimental, sentiment*).

The most complex entry of the set, also the basal entry in word-formation terms is:

> **sentiment** *n* **1 a**: an attitude, thought, or judgement prompted by feeling: PREDILECTION **b**: a specific view or notion: OPINION **2 a**: EMOTION **b**: refined feeling: delicate sensibility esp. as expressed in a work of art **c**: emotional idealism **d**: a romantic or nostalgic feeling verging on sentimentality **3 a**: an idea colored by emotion **b**: the emotional significance of a passage or expression as distinguished from its verbal context

2.2. Disregarding 1a and 1b as semantically irrelevant, and reducing the remaining definitions (glosses) offered by this entry to *key words* (i. e. excluding structural items and explanations) we arrive at the following inventory:

First-order key words	Second-order key words (Qualifiers)
feeling (2)	*refined*
emotion	*delicate*
sensibility	*emotional*
idealism*	*romantic*
sentimentality*	*nostalgic*

* Sentimentality will be excluded from further analysis as leading to a circular definition. *Idealism* will also have to be disqualified as restricted by *emotional* (cf. the entry above).

Applying the same procedure to all the five "full" MW9 entries (containing definitions) that center around *sentimentality*, we expand the above inventory to:

First-order key words

feeling (5), *sensibility* (3), (*sentimentality*, 3), *sentiment* (3), *emotion* (1), (*idealism*, 1), *affectation* (1)

Second-order key words (qualifiers)

emotional (3), (sentimental, 3)**, refined (1), *delicate* (1), *romantic* (1), *nostalgic* (1)

** Excluded from further procedures as leading to circular definitions.

131

2.3. We can now use this expanded key-word list to go back to MW9. There, we look up all these words and further expand our key-word list from the definitions containing at least one of the key words already listed, and disregarding definitions without them. We may call this the "principle of intersection as internal corroboration".

Here is how it works on the example of the entry:

emotionalize *vi* emote, give free play to emotions, make a scene

Both *emote* and *(give free play to) emotions* qualify as new key words/phrases, while *(make a) scene* is eliminated as it is not corroborated by any earlier occurrence of *scene*. Most human analysts could, of course, be relied to disqualify intuitively the item of *(make a) scene* as a lexical expression of the concept of "sentimentality". Here, however, we have lexicographic material self-expanding and self-controlling toward a concept in a process of the lexicalization of that concept. A computer can readily be imagined executing this entire procedure, especially since all leading dictionaries nowadays are produced and printed by computers.

Here is how the expanded key-word list originating from the entry *sentiment* further expands itself through lookups in the MW9 and intersections described in the preceding passages:

> **feeling** *n* yields *sensitivity, susceptibility*
> **feeling** *adj* yields *sentient, sensitive, easily moved, sensitivity* (also the run-on entries *feelingly, feelingness*) ·
> **sensibility** *n* yields *sensitiveness* (also the run-on entries *sensible, sensitive* adj, *sensitive* n, *sensitively, sensitiveness*, sensitivity)
> **emotion** *n* yields *excitement, affective*
> **emotional** *adj* yields nothing (but provides the following run-on entries: *emote, emotionality, emotionalism, emotionalist, emotionalistic, emotionalize*)
> **refined** *adj* yields nothing
> **delicate** *adj* yields nothing (except the run-on entry *delicacy*)
> **romantic** *adj* yields *romance, lovemaking* (plus the run-on entries *romantic* n, *romanticism, romanticize*)
> **nostalgic** *adj* yields nothing (but the run-on entries *nostalgia, nostalgist*)

2.4. The new, now twice-expanded, list of key words can now be used again to turn back to the MW9, making it "feed upon itself" until it comes up with a third list, and so forth. The question, however, is: when do we reach the point of diminishing returns? Deciding about when to stop all this lexicographic recycling could be helped by reordering the list of key words as it grows according to the incidence of intersection of each key word. This would clearly indicate which key words are the most central to the lexicalized concept of "sentimentality". Whatever we did, we would in all probability soon run out of steam (read: out of new key words), exactly because of the requirement of corroborative intersection built into the method, but also ultimately because every dictionary is a finite inventory of items.

132

3.

3.1. Therefore, it we need more vocabulary items, to further lexicalize the concept of "sentimentality", we should look to a more elaborate repository of lexicalized concepts: a *thesaurus. Roget's International Thesaurus* (4th ed.),[4] one of the best-known thesauri of English, contains over 250,000 items in 1,042 categories/fields. Looking for "sentimentality" entails going through a number of steps. First, out of the 8 fundamental *classes* (Abstract Relations, Space Physics, Matter, Sensation, Intellect, Volition, Affections) we select the last listed. It breaks down into 4 *subclasses*: Personal Affections, Sympathetic Affections, Morality, Religion. The first of them offers Emotion (one of 9 such *divisions*), finally taking us to Feelings,one of two *fields* that Emotion is made up of (the other being Lack of Feelings). The long search is not quite over: Feelings (listed as 855th field in the Thesaurus) contains a total of some 300 single items and expressions arranged into 10 nominal, 8 verbal, 8 adjectival and 3 adverbial paragraphs.

3.2. The simplest approach now is to look for a paragraph actually entitled "sentimentality".The paragraph in question is numbered 8 and this is how it looks (boldface indicating the key items):

> **sentimentality, sentiment, sentimentalism,** oversentimentality, oversentimentalism, bathos; nostalgia, nostomania; romanticism; sweetness and light, hearts-and-flowers; bleeding heart; mawkishness, cloyingness, maudlinness, namby-pamby, namby-pambyness, namby-pambyism; mushiness *or* sloppiness [both informal]; **mush** *or* slush *or* slop *or* goo *or* schmaltz [all slang]; sob story *or* tear-jerker [both slang], soap opera

The slant toward "oversentimentality" is obvious. We therefore look next for items more central to the concept of "sentimentality" among other nominal paragraphs. We do this by using the most frequent key words obtained from the MW9: *feeling, sensibility, sentiment,* and they direct us to paragraphs 3 and 4. These, we discover, are slanted toward "emotion", "sensation" (3) and "sensitivity", "delicacy" (4). A slight morphological shift toward the plural form steers us to paragraph 1 that turns out to be the closest so far to what we intuitively feel the concept of "sentimentality" contains:

> **feelings, emotions, affections, sentiments, passions, sensibilities, susceptibilities, sympathies,** tender susceptibilities, finer feelings; the logic of the heart; emotional life; affectivity, affective faculty; feeling tone

4. Revised by Robert L. Chapman (Crowell, New York, 1977).

3.3. Our lengthening list of lexical items expressing the concept of "sentimentality", however indirectly and marginally with some items, could be expanded further by several procedures. We could make the Thesaurus "feed upon itself", that is look for repeated occurrences of the items already on the list within the field 855. Or we could carry out this search by combing the entire Thesaurus for other occurrences of such items. We might refine the search by using the Thesaurus index with its copious numerical cross references to other fields and paragraphs. All of these procedures, however, would allow some useful items like *corny* (883.9) and *hokey* (22.8) to slip through the net. This simply because their microcontexts (set off by semicolons) contain no item from our list of lexicalizations of "sentimentality", so no intersection would take place.

We should note that searching the MW9 (cf. 2.4), *would* have turned up both *corny* and *hokey* since both their definitions contain the item *sentimental* (and *mawkish*).

> **corny** (sense 3) mawkishly old-fashioned: tiresomely simple and sentimental (see **trite**)
>
> **hokey** (sense 1) sickly or affectedly sentimental (see **corny, mawkish**)

3.4. An alphabetized thesaurus is easier to search, and if such a thesaurus is considerably larger than the Roget's International, the results — the list of items lexicalizing "sentimentality" — is certain to be more spectacular. Rodale's *Synonym Finder*[5] is such a thesaurus and, containing over one million synonyms, it is four times the size of the Roget's International. Let me illustrate this with the entries *sentimentality* and *sentimental* from the Synonym Finder:

> **Sentimentality,** *n* emotionalism, gushiness, mawkishness, maudlinism, maudlinness, sentimentalism; emotiveness, emotivity, play on *or* appeal to the emotions, melodrama, melodramatics, dramatics, theatrics, playacting; corniness, hokiness, affectation; bathos, mush, *Inf.* glop, treacle, *Sl.* hearts and flowers, *Sl.* schmaltz, *Sl.* goo; *Sl.* sob stuff, *Sl.* sob story.
>
> **Sentimental,** *adj* **1.** emotional, feeling, warm, tender, affectionate, loving, soft-hearted, warm-hearted, sympathetic, compassionate; impressionable, sensitive, receptive, responsive. **2.** maudlin, melodramatic, mawkish, misty-eyed, dewy-eyed, gushy, *Austral.* soony, *Inf.* drippy, *Sl.* icky; soupy, romantic, with the soul of a shopgirl, *Inf.* mushy, *Inf.* slushy, *Inf.* gloppy, *Sl.* schmaltzy, cow-eyed, salt-eyed; teary, lachrymose, weepy, *Inf.* drippy, choked up, soppy; sloppy, uncontrolled; babyish, namby-pamby, insipid; over-romantic, *Inf.* corny, *Sl.* hokey; affected, *Inf.*, bleeding heart; nostalgic, pining, languishing.

5. By Jerome Irving Rodale (rev. by Laurence Urdang and Nancy LaRoche), Rodale Press, Emmaus, Pa., 1978.

Whether we now utilize the key-word list obtained and expanded from the MW9, or use the Synonym Finder's contents to feed upon themselves — even though either procedure is limited by the requirement of intersection — the yield can be anticipated as quite impressive. How can it be anything else when entries the size of *feeling(s)* (77 items), *sensibility* (25), *emotional* (63) and *emotionalism* (33) are involved?

4.

4.1. But what is the point of all these exercises? What is the use of these lists (varying in length and structure) of lexical items associated with or emanating from the central concept of "sentimentality"? A more complete awareness of the extent and variety of the lexicalization of this concept is by itself a sufficient justification for the effort.

These lists, however, have another purpose to serve — more valuable and more exciting in sholarly terms. They can also be used as an efficient searching tool *to discover the presence of the concept of "sentimentality" diffused in text.* To the degree that words express concepts — to go back to the initial reasoning of this paper — words in a text can be used as the evidence of presence in the text of specific concepts expressible by these words. What is more, an unanticipated, unlikely or totally unpredictable presence of such lexical items (indicative of specific concepts) in certain contexts *may contribute toward a redefinition of the concepts themselves.*

4.2. "Sentimentality" is a tipical diffuse concept that can be expected to benefit from this approach, from such lexical searches of text. Modern technology makes it possible now to search quickly and reliably texts of unlimited size. A single laser-read compact disc can now contain texts totaling a mind-boggling 250 million words, or nearly six entire Encyclopaedia Britannicas. Well, the Britannica (with its over 42 million words) is exactly the kind of text that may be profitably searched for the lexical presence of such a culturally diffuse concept as "sentimentality". To prove my point, I offer here a modest simulation of what a computer might do.

4.3. Told to look for the word *sentimental* and its derivatives (*sentimentally, sentimentality* etc.) in the Britannica, it would find them, of course, in such more or less predictable entries of the Micropaedia[6] as:

> **Foster, Stephen Collins** (vol. IV) "...composer whose popular minstrel songs and *sentimental* ballads..."
> **Disney, Walt** (vol. III) "'Snow White' was widely acclaimed by critics and audiences alike as an amusing and *sentimental* romance"

6. *All references are to the New Encyclopaedia Britannica (in 30 vol.) of 1974.*

operetta (vol. VII) "A musical—dramatic production similar in structure to an opera but characteristically having a romantically *sentimental* plot, interspersed with songs ..."

(and later in the text)

"... influenced by the gentler quality of Viennese operetta the French style became itself more *sentimental* and less satirical"

Biedermeier style (vol. II) "... Biedermeier painting, which was either genre or historical and most often *sentimentally* treated."

More marginal lexical items can also lead to areas of undoubted presence of the concept analyzed ("sentimentality"), as illustrated by *pathos* and *nostalgic* — both from field 855 in the International Thesaurus — as they turn up in the more general, long articles of the Macropaedia:

Popular Music (vol. 14, p. 811 /"Country music"/) "... a range of songs of commercialized *pathos* treating of dying mothers, orphaned children, bereft lovers, and lonely men far from home. During World War II 'country music' expanded enormously ... probably because the strong *nostalgic* bias of the music was particularly appealing during a time of ..."

Dickens, Charles (vol. 5, pp. 707, 708) "...'Pickwick' displayed ... many of the features ... throughout his fiction: attacks ... on social evils ... *pathos*; a vein of the macabre ..."

"... in 'Oliver Twist' — the spectacle of the lost or oppressed child as an occasion for *pathos* and social criticism."

"... the death of Little Nell ... a few decades later it became a byword for 'Victorian *sentimentality*'."

"In Paul's ['Dombey and Son'] early death Dickens offered another famous *pathetic* episode;"

4.4. It should not be too hard, from the above short list of examples located intuitively, to imagine literally hundreds of such occurrences dispersed throughout the Encyclopeadia Britannica. Offered, in the form of a computer printout, as an integral corpus, easy to consult and immediately suggestive of possible synthetic procedures, these selective keyword-in-context concordances must be seen as a promising and exciting new tool of conceptual research.

LEKSIKALIZIRANJE POJMA "SENTIMENTALNOST"

Pretpostavka da "riječi izražavaju ideje" provjerava se najprije na čestotnoj listi generalnog vokabulara američkog engleskog. Pokazuje se da se riječi koje izražavaju najčešće pojmove s liste (čovjek, vrijeme, prostor) poklapaju s najčešćim riječima na sličnoj čestotnoj listi za ruski jezik. Time je dokazana kako univerzalnost najčešćih pojmova tako i valjanost polazne pretpostavke. U radu se zatim ispituje kako se pojam "sentimentalnosti" leksikalizira. Kao prvo, izvlače se ključne riječi iz definicija očito relevantnih natuknica (koje sadrže osnovu *sentiment-*) u jednom tzv. akademskom rječniku engleskog jezika. Te ključne riječi dovode do drugih natuknica s njihovim definicijama, čime dobivamo nove ključne riječi. Ova proizvodba leksičkih jedinica koje izražavaju pojam "sentimentalnost" znatno se proširuje u sljedećoj fazi, kad se dva velika tezaurusa engleskog jezika podvrgavaju sličnom postupku. Na kraju se uvjerljivo pokazuje da su tako dobivene liste leksičkih jedinica vrlo korisno sredstvo za (ovdje simulirano) kompjutersko određivanje prisutnosti, posredne i neposredne, difuznog pojma "sentimentalnost" u tako ogromnoj i raznolikoj građi kakav je tekst Enciklopedije Britanike.