

USAGE IN THE *USAGE DICTIONARY OF ANGLICISMS*
IN SELECTED EUROPEAN LANGUAGES

MANFRED GÖRLACH

University of Cologne

Multa renascentur quæ iam cecidere cadentque
quæ nunc sunt in honore vocabula si volet usus
quem penes arbitrium est et vis et norma loquendi.

(Horace, *Ars Poetica*)

Many a word yfalne shall eft arise
And such as now bene held in hiest prise
Will fall as fast, when vse and custome will
Onely vmpiers of speach, for force and skill.

(Puttenham's translation, 1589)

1. Introduction

All through the history of dictionary-making and reflections on standard languages, usage has been one of the lexicographers' major stumbling-blocks.¹ Although it can be argued that the question of a linguistic norm is even more relevant in the fields of phonology and syntax, it is eminently important for vocabulary, too. To indicate the currency, degrees of acceptability and the geographical, social and stylistic restrictions of individual words is a problem common to both more prescriptive and more descriptive dictionaries. Johnson's aim to establish a standard for the English lexicon left but a few words unaffected in his great dictionary of 1755 and these had usage labels attached to such entries to warn readers against using them; modern dictionary editors sometimes

¹ This paper was improved by various critical readers; I wish to thank in particular Helen Weiss, John Davis and Keith Battarbee. For definitions and practical applications of 'usage' cf. *inter alia* Fowler (1965) and Todd – Hancock (1986). It is significant that the term is first recorded in a linguistic sense from Defoe, 1697, "referring to the proposed English Academy to monitor the language: 'The voice of this society should be sufficient authority for the usage of words' " (quoted from McArthur 1992: 1071).

use panels of experts to decide on the acceptability of words and their individual meanings. Whatever method for establishing usage is preferred, whether that of the editor's *Sprachgefühl* or the combined judgment of experts (a method which appears to permit neatly graded degrees of acceptability), any indication of usage carries with it idiosyncratic elements as well as the restriction to a particular period, region and class-related opinion. It would be silly to complain about this, since it is the very essence of usage, that the labels used to indicate it have a certain degree of vagueness about them. The narrower the description of correct usage is, the more certain it is that the evidence will be homogeneous, and the more likely that the guidance will be taken as prescriptive.

In a dictionary which has 'usage' as the first word of its title, readers will rightly expect this type of information to be central. The *UDASEL* project was in fact started with the explicit intention of concentrating on the fully *accepted* loanwords from English in 16 European languages,² rather than focussing on more narrowly etymological questions of word origins. As will be obvious (or become evident from my discussion), the difficulty of describing the usage of specific items in an individual language multiply when 16 sets of data are compared (which makes possible up to 240 binary comparisons).

The basic decision as far as the currency of loanwords from English is concerned was not to rely on text corpora, for various reasons:

- a) such collections from recent texts were not available for most of the languages concerned;
- b) where such corpora existed, their representativeness and comparability was open to doubt;
- c) the frequencies elicited from them would not really be indicators of acceptability: for this the cotext and context would have to be investigated for each individual item, and judgments would therefore in any case be based, at least partially, on the interpreter's *Sprachgefühl*.

All this made it necessary, if the evidence was to be covered in a 'snapshot' manner for the early 1990s, to rely on the competence of educated users of the individual language, and leave judgments on the acceptability of English words to the specialist. This hard-won decision is supported by the fact that pilot studies with German informants (Görlach 1994) have shown that speakers' views on degrees of acceptability conform to a greater extent than might have been expected, or feared.³

² For further information see Görlach (1994, forthcoming a, forthcoming b); words are included only if they have something English in their form at least in one of the 16 languages – contrast the different principles in other dictionaries of anglicisms, such as Carstensen (1993-1995).

³ Whether the informants have a command of English is certainly relevant in many ways; however, it is not the case that bilinguals are more lenient in accepting English words as part of the German language. Educated speakers if anything tend to be more purist, i.e. they make greater efforts to keep their languages apart.

2. Degrees of acceptability

After the pilot study was completed, the grading decided on was based on a combination of degrees of integration and of acceptability/currency: these two aspects are not always compatible, but coincide in most cases, and they were combined in order to prevent the system from becoming too complex for a general reader to take in and remember. The stages reflecting degrees of integration plus currency are, then, divided into three; they are here listed on the basis of the advice sent out to colleagues collaborating on the project, and exemplified by typical cases.

2.1. Words not forming part of the language in question

The item in question is not felt to be part of the individual language; subcategories are marked as follows:

- : the word is not known
- 0 : the word is known only as a quotation word (in code-switching) and is only understood by bilinguals
- Ø : the word is known and used only with reference to foreign objects (i.e. it is a foreignism).

When does a word become a component of *la langue*? Strictly, the question is equally impossible to answer both for newly coined *ad hoc* compounds and derivations and for foreign words. It is comparatively easy to say that a word is completely unknown (though it will not be to people knowing English!) – but the distinction as applied to words that do occur in, say, newspaper advertising but are not meant to be part of the receiving language is gradual, for the categorization 'not part of the language' is somewhat subjective. Here are a few illustrations from my native German: we make no use whatsoever of items like *after all*, *after-skiing**, *all hands*, *armwrestling* or *badlands* (some* are not even English!), whereas we might find (but not accept as German) items like *all right* (contrast the complete integration of *o.k.*), *barman* or *beauty* – but are *barbecue*, *brunch* and *bike* on their way in? Is *big mac* becoming generic for *junk food* (and is this a new term in German)?

Words relating to the foreign culture (and occurring as technical terms or for local colouring) are 'foreignisms'. Thus, we are likely to come across items like *acre*, *alderman*, *batsman* or *bed & breakfast* in descriptions of Anglo-Saxon culture, but they have no proper referents in German culture. (Again, facetious uses in advertising etc. can present a problem, as when pubs begin to be thus called in names, while the word is never found generically – the *UDASEL* sometimes indicates '1w' = 'found in written form only', but this is not fully satisfactory.)

All these 'zero' items should be omitted, then, from monolingual dictionaries of German (but some dictionaries are very generous in admitting marginal words and also include names). However, zero categories become relevant if, as frequently happens, such words *are* found in a few languages, and their absence from or marginality in the others must then be indicated in a contrastive dictionary, such comparisons being the central aim of the *UDASEL*. Thus, many languages have *weekend* as a regular item (French, Russian, or Hungarian), whereas in Dutch or Italian the word competes with native equivalents, and in German and Finnish native words only are used. It may happen – and the fact often comes as a surprise – that no other language has a particular loanword well known from one's own language – like *cleverness*, *flummery* and *gully* in German.

2.2. Restricted currency

If a word is accepted, i.e. is considered to be part of their language by educated native speakers, it may be restricted or common;⁴ restricted use often goes together with incomplete integration. The restrictions (with their causes and consequences) may be of very different types which can combine in many complex ways – or may differ for different sections of the speech community. Usage may be:

- a) related to age/time – (1a) for 'archaic', (1o) for 'obsolescent',⁵ (1m) for 'modern/fashionable'. Archaic terms may have become obsolete because of cultural change, or been edged out by purist measures, or may have been replaced through changes in usage which are difficult or impossible to determine. Many such words are found in older editions of dictionaries,⁶ or remembered by older speakers – an inevitable mirror image of (1m) 'incoming fashionable lexis which is not certain to last'⁷ or (1y)

⁴ The much-quoted graph in the *Shorter Oxford Dictionary* which distinguishes between the common core and various fields of restrictive use can serve as an example; in lexicographic practice common words are unmarked, restricted usage is labelled.

⁵ Such words are comparatively easy to document where early dictionaries are available; there is even a substantial loss of early anglicisms (accordingly marked '0a') as Liebknecht (1912) and Dunger (1909) document for German.

⁶ Since dictionaries tend to mark restrictions insufficiently or not at all, it is often impossible to reconstruct usage values at the time of publication, but since the aim of the *UDASEL* is to indicate synchronic currency, anglicisms mentioned in early publications can be tested as to their present-day currency: do they survive at all, and are they restricted to archaic use, or to designate historical phenomena, or have they retained their full currency? Accordingly, labels like '1a' or '1o' can be attributed on the basis of present-day competence. Note that a few early dictionaries at least distinguish between '0' and '1', as Liebknecht (1912) did by using different fonts in headwords.

⁷ Again, the difficult decision on when a word becomes 'dictionary-worthy' is not confined to the *UDASEL*. Currency may well be overridden by 'historical' interest – the fact that a word is current in newspapers in a certain period justifies its inclusion as forming part of cultural history. Such is the case with the American import *airbagging* which appears to be restricted to 1994 Germany – more or less (Görlach:

'characteristic of the language of younger speakers'. Changes of usage can make items seemingly contradictory, both (1a) and (1m), strange as it may seem. Thus, anglicisms in football and tennis terminology were replaced by German equivalents in the early 20th century, but modern journalese often prefers (or uses as an alternative) the English words (*Torwart*: goalkeeper, *Schiedsrichter*: referee etc.);

- b) (obviously) related to the specificity of designation; words referring to the specialized entities of science and technology, (pop) music and drugs obviously do not form part of the common lexis and should be marked (1t).⁸ Obviously, interest in, and knowledge of, special fields can change; highly technical tennis terms have become relatively well-known, but this does not apply to those of golf; computer terminology is becoming more common, whereas that of the traditional crafts has been lost, etc. Technical specialization combined with highly colloquial features of 'anti-language' produces the highly pertinent category of 'slang' (1s);
- c) distinguished stylistically in various ways. The scale of formality is especially important where 'colloquial use' (1c) is a characteristic feature of an anglicism – such style values are often correlated with the written vs. spoken dichotomy. Restrictions to domains and text types are of obvious relevance in fields like media language ('1j' for 'journalese', including 'advertising'). Another obvious domain is much less affected by English influence – literature (1b for 'belles lettres');
- d) related to evaluations, which often explain why one word is preferred to another. An English word may be used in a derogatory (1d), euphemistic (1e), facetious (1f) or pejorative (1p) way – the easy availability of English to most members of the speech community can be utilized and words adopted to bring out nice distinctions of attitude and expressiveness;⁹
- e) the result of regional differences causing different degrees of acceptability. This is quite clear where national standards are concerned – anglicisms do not have the same currency throughout France, Belgium,

forthcoming a). Since each of our collaborators (mostly middle-aged or older) speaks a 'chronolect' it has proved immensely useful and necessary to involve younger speakers, like research students, to get our information on the existence of words and their currency right, as far as the comprehensive usage of all age groups is concerned.

⁸ This should exclude words so technical and specialized that they are unlikely ever to occur in a newspaper context – but since all collaborators were asked to respond for their language to items suggested for another language, the entire collection has drifted more to the specialized end of the lexis than was initially intended. (Also, some collaborators appear to have included more technical words from the beginning, including words relating to historical objects – such as obsolete items from ship-building.)

⁹ It is no surprise that a few four-letter words (like *shit!*) have a certain currency because they sound less offensive than native equivalents – English, too, has such uses for tabooed domains from other languages.

French-speaking Switzerland and Luxemburg, and the differences between Austria, Switzerland and Germany (former East vs. West) are often glaring;¹⁰

- f) affected by normative measures influencing the distribution of loanwords, e.g. recent French legislation (1x = banned), but also by less apparent linguistically puristic trends.¹¹ (It is an open question whether a greater reluctance to adopt anglicisms can be ascribed to structural differences in the receiving languages – which in Europe belong to quite different types – as is illustrated by the ‘purist’ attitudes and their effects in Iceland, France, Finland or Greece);
- g) restricted to rare occurrence in the case of some items for any of the above reasons, but a low token count can also be due to more general factors, such as the availability of less marked synonyms, misleading connotations, danger of misunderstandings – or users’ problems with correct pronunciation or spelling. In all these cases, we thought it better simply to note the rarity of an item (1i – infrequent).

Restrictive labels can be combined, especially where they relate to different classes of factors from a)-g) above. Thus (1ymi) is a rare modernism among young people, (1tro) a technical term of regional currency, now going out of use.

Every linguist will have qualms in settling for a fixed usage value for an individual word – or for its several meanings. And yet, there is a kind of common consent which can be established among educated native speakers (the *consensus eruditorum*). How well this can work in international cross-linguistic comparison is a different matter. Comparisons are the more effective the less detailed the classifications and descriptions of national usages are: the more specific the data on currency, acceptability, denotations and connotations in highly divergent societies, the more meaningless comparisons will become.

2.3. Fully accepted items

An item’s progression to the ‘fully accepted, but still recognizably English’ category (2) is, again, riddled with difficulties. How far can we be certain

¹⁰ As far as anglicisms in German are concerned, the most obvious differences are in the fields of sports (mostly football and tennis), but gallicisms survive in the domains of the railway, the postal services and in many fields of daily life (in Switzerland also due to continued language contact). The linguistic purges of the East German authorities, 1949-89, proved less effective than expected (Lehnert 1990).

¹¹ The topic deserves a thorough contrastive study. Normative measures frequently present only the tip of the proverbial iceberg – French language laws for instance affect the use of technical terms but leave the ubiquitous *snackbars* and *hotdogs* of Paris unscathed (cf. Görlach: forthcoming c). Frequently, purism affects only the formal style (anglicisms in Icelandic, gallicisms in 19th-century German) or is restricted to certain regions (measures against gallicisms in Flanders, but not in the Netherlands; against anglicisms in early 20th-century Germany, but not so much in Austria and Switzerland, etc., cf. e) above).

about the general currency of a term? Does it mean by definition that ‘2’ words do not have, and cannot have, restrictive labels? Practice has shown that the sequence 0-1-2 can be used, and we think should be used, to indicate currency somewhat more loosely than initially defined, which makes combinations other than ‘1t’ meaningful. Thus ‘0t’ can be used for a technical term that a non-specialist newspaper reader is not likely to guess even when met in its proper context, ‘2t’ the usage value of a word which retains its technical reference though it has become more widely known (at least passively), as some computer terms have in recent years.

The distinction between ‘2’ and ‘3’ is one of form – ‘3’ words being no longer recognized as English in spelling, pronunciation and morphology, or in at least one of these. The criteria appear to be unambiguous, but their application poses new problems. The integration of a foreign word can proceed at varying speeds on different levels; the process will be determined by both the linguistic structure of the receiving language and the attitudes of its users. Thus, many loanwords present no phonological/graphemic problems at all, for example *bit*, *bitter*, *box*, *test* etc. in German – and it might therefore be asked whether the 2/3 distinction is useful. However, the plurals *Boxen* vs. *Tests* provide at least a marginal criterion for calling *Box* (3) and *Test* (2), unless we wish to see the general currency of *Test* as sufficient for its classification as (3) – again extending the use of the 0-1-2-3 sequence to indicate currency and acceptability.¹²

The distinction 1 and 2/3 breaks down, however, in languages using a non-Roman alphabet and adopting loanwords in their phonemic shape. In Russian and Bulgarian nearly all items have to be marked ‘3’ regardless of their currency – which may be very restricted. ‘3ti’ thus is to be interpreted as equivalent to ‘1t’ in other languages and is represented thus in the grids – otherwise these two eastern languages might misleadingly show full acceptability where western languages might indicate restrictions.

3. Statistical analysis

How far can we base any statistical comparisons on classifications which are admittedly impressionistic? If we are willing to disregard the fact that ‘current/accepted’ may well mean different things for different speech communities (and individual collaborators), we can check the occurrence of selected words and arrive at a sum which will hopefully tell us something about the openness of a particular community/language to English lexical influence. I have here

¹² Two further numerals used are not related to usage and currency: ‘4’ is for native words which happen to be the same as in English (often for reasons of genetic relationship) and ‘5’ for words that come from other, non-English languages, such as the large number of Romance words which, if borrowed from English into Romance languages, mostly ‘return’ to their Romance pronunciation, and, not being marked as English in their form, are excluded from the dictionary.

selected for words beginning in A-G the 278 items which are widespread enough to justify the use of a grid to represent their geographical distribution, and have distinguished between absence (-, 0, Ø), restricted currency (1) and acceptance (2, 3). Some problems remain with 3-5: some items *are* integrated (and languages using non-Roman alphabets tend to 'integrate' by the very procedure of transliteration) – '3t' is here best equated with '1t' as indicating restricted currency (as '3a' etc. do). '4' must necessarily count as 'absence' since no overt anglicism is found, as must '5' where languages have a related word from native stock or another language (often Latin). I am aware of problems arising from this kind of count, but cannot think of any better solution.

A further problem is well-known in all disciplines that have to rely on the judgment of fieldworkers. Wherever distinctions are not clear-cut they have to be made by the researcher and individual values allocated to the item in question. Such judgments will depend on features of personal experience as well as the linguistic situation.¹³ Looking at the evidence collected for the *UDASEL* entries so far, one gains the impression that the specialists responsible for Dutch, Italian, and Polish are more ready to assign a '2' (for 'fully accepted') than those dealing with Icelandic, German, French, Spanish, and Russian, who prefer a '1' (for 'restricted usage'), an impression that can be verified from the statistics for the 278 selected items. Are these differences 'real' (i.e. based on hard linguistic evidence), or do they reflect the personalities or culturally-influenced perceptions of the researchers? There is no objective way of testing this; even a replication of the work – if it were feasible – would not be a solution, since the conditions would not be the same.

The following differences are striking (but not all are easy to explain): whereas all languages have more items marked as restricted than as fully accepted, except for Finnish,¹⁴ the sequence from a larger to a smaller number of '2' markings, giving the percentage of numbers in categories 2:1, is as follows: Dutch 85, Norwegian 80, Polish and Rumanian 70, Croatian 67, Hungarian 64, Italian 58, German 52, Icelandic 39, French 38, Spanish 24, Albanian 12. Languages with non-Roman alphabets might be expected to have integrated items throughout as a consequence of transliteration; however, the practice of labelling shows that the collaborators have (rightly) based their judgments on criteria like currency, morphological adaptation and attitudes as well. The differences between Greek 65, Bulgarian 43 and Russian 35 can therefore be interpreted, with

¹³ The problem of 'iso-fieldworkers' is well-known in dialectology. The term refers, not entirely seriously, to lines cutting across linguistic continua not because linguistic features differ on the two sides of the line, but because the methods or the qualifications of the fieldworkers investigating the two areas were different.

¹⁴ The figures have no absolute value since the selection of the grids was made to show significant regional distinctions; in consequence, many widespread and fully acceptable items were left out. Finnish appears to have more integrated loans because, for the relative small number of loanwords it takes in, morphological adaptation is (still) the rule.

due caution, as reflecting the assimilatory power of these languages.

For these and other reasons arising from the social, geographical and stylistic heterogeneity even of modern standard languages, statistics based on the *UDASEL* evidence must be taken with a certain degree of caution; however, the data are the best available and, I would like to argue, the best that it is possible and feasible to obtain. The map and graph below are based on 278 items exhibiting 'interesting' distributions which qualified them for visual representation in the form of grids. These are arranged so as to be as geographically iconic as possible: the Germanic languages in the NW, and again Icelandic and Norwegian in the north, Dutch and German in the south, and so on. Even regional distributions such as 'Balkan' can be seen at a glance. Usage is here reduced to black for 'absent', hazed for 'restricted' and white for 'accepted'. A few specimen grids will help to illustrate the principle and the type of information that can be drawn from the patterns. (In the *UDASEL* these will of course be accompanied by very detailed comment and language-specific data.)

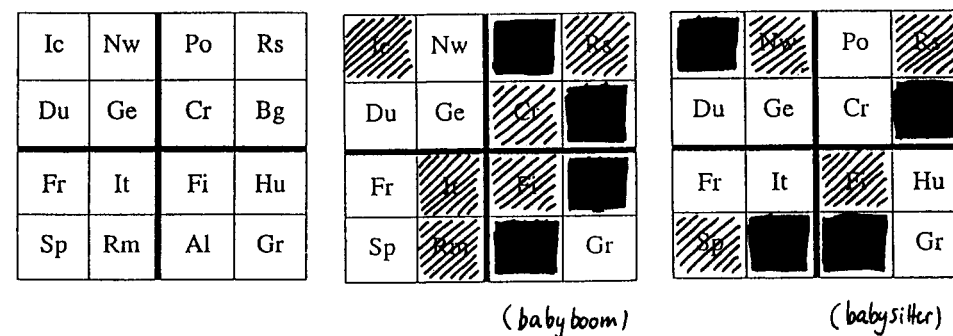


Figure 1

Counting the labels for the 278 grids envisaged for letters A to G, we arrive at the figures given in the Fig. 2.

To what extent can we properly base sociohistorical conclusions on this type of evidence? It is quite clear that analysts will have to turn to more limited sets of data, comparing the evidence from within, say, the Slavic languages, or looking at the timescale of borrowings, or correlating linguistic structure and attitudes gleaned from other sources with the data here made available for such contrastive investigations for the first time. In a field in which every word has its own history, statistical comparisons are necessary for the general information they provide, but they certainly need to be complemented by in-depth studies of more specialized aspects. The *UDASEL* will now provide a framework into which the findings of such microstudies can be fitted.

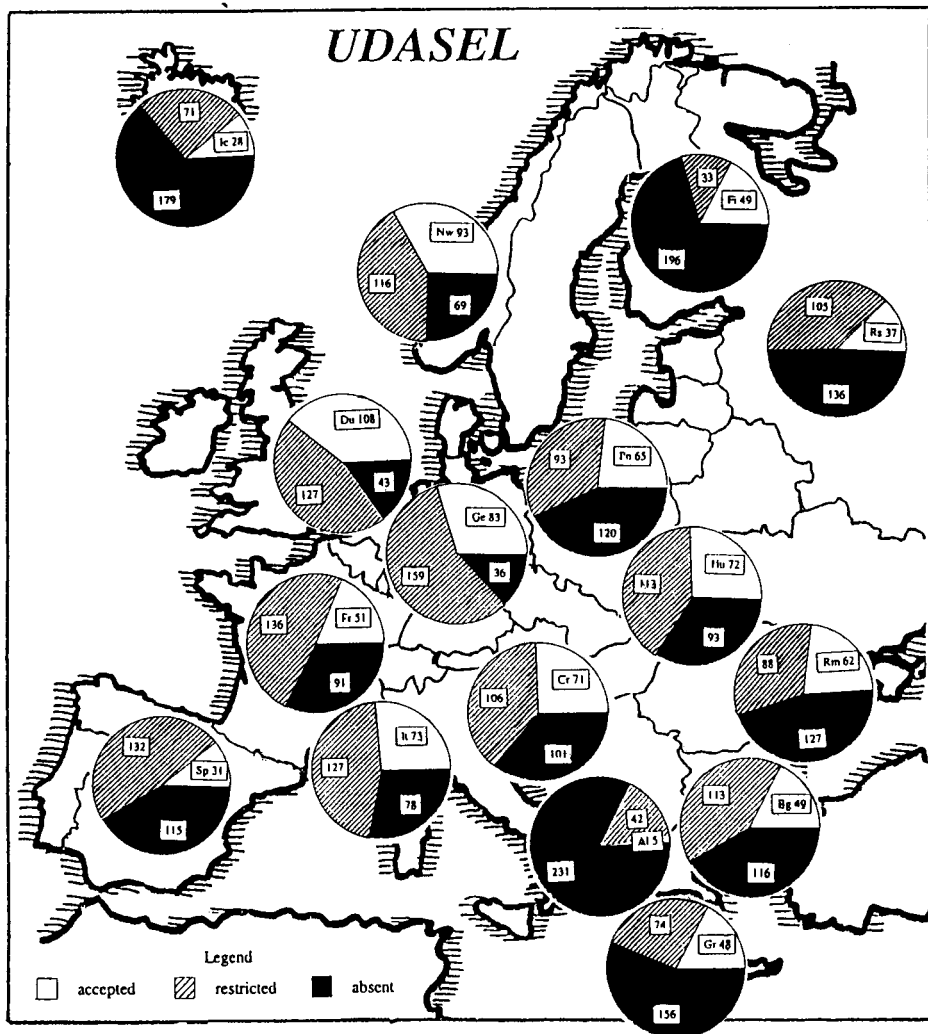


Figure 2. Currency: Provisional results based on 278 grids (letter A-G)

REFERENCES

- Carstensen, Broder
1993-1995 *Anglizismenwörterbuch*. (ed. Ulrich Busse.) 3 vols. Berlin, New York: Mouton de Gruyter.
- Dunger, Hermann
1909 *Engländerei in der deutschen Sprache*. Hildesheim: Olms.
- Fowler, Henry W.
1965 *A dictionary of modern English usage*. (2nd edition.) London: University Press.
- Görlach, Manfred
1994 "A Usage dictionary of anglicisms in selected European languages", *International Journal of Lexicography* 7: 223-246.

- forthcoming "Is airbagging hip or mega-out? A new dictionary of anglicisms", in: Martin Pütz a (ed.).
- forthcoming "The Usage dictionary of anglicisms in selected European languages: Progress, problems and prospects", in: *Proceedings of the Cracow Conference of March 1996*.
- forthcoming "Purism and the UDASEL", in: *Festschrift for Rudolf Filipović*. Zagreb: Institute of Linguistics.
- Görlach, Manfred (ed.)
forthcoming *The UDASEL. 3 vols. (The Dictionary. An annotated bibliography. English in Europe – a volume of national studies)*. Oxford: OUP.
- Lehnert, Martin
1990 *Anglo-Amerikanisches im Sprachgebrauch der DDR*. Berlin: Akademie-Verlag.
- Liebknrecht, Wilhelm
1912 *Volksfremdwörterbuch*. (13th edition.) Stuttgart: Dietz.
- McArthur, Tom (ed.)
1992 *The Oxford companion to the English language*. Oxford: OUP.
- Pütz, Martin (ed.)
forthcoming *Language choices: Conditions, constraints, and consequences*. Amsterdam: Benjamins.
- Todd, Loreto – Ian Hancock
1986 *International English usage*. London: Croom Helm.
- Viereck, Wolfgang – Wolf-Dietrich Bald (eds.)
1986 *English in contact with other languages*. Budapest: Akadémiai Kiadó.