GERMAN LOAN WORDS AND CULTURAL STEREOTYPES

Michael Stubbs

An analysis of loanwords from German into English, found through computer-assisted lexical research.

The originators of the Oxford English Dictionary were in no doubt about the cultural importance of providing an exhaustive record "of the appearance of each word in the language". It was argued that words "embody facts of history", and are "a record of great social revolutions, revolutions in nations and in the feelings of nations". Much may be learned, it was argued, "by noting the words which nations have been obliged to borrow from other nations" (Trench 1851, 1858). And words are often created or borrowed into a language in response to world political events: cold war 1945, sputnik 1957, Watergate 1972, perestroika 1987, intifada 1988. Dates are the first attested uses in English.

WORDS, POLITICS AND NATIONAL STEREOTYPES

If English speakers are asked to list the words that English has borrowed from German, then they may well think first of a set of loans from before and during World War 2, such as Nazi 1930, Third Reich 1933, Nazism 1934, Führer 1934, Gestapo 1934, Luftwaffe 1935, Anschluss 1938, and Blitzkrieg 1939.

Quite apart from such historically significant cases, loan words often confirm national stereotypes, and symbolize the foreign and the strange. From German, English has borrowed blitzkrieg, kitsch and lederhosen. From French, bistro, gigolo and mot juste. However, a more systematic search for German loans in English does something to balance such crude national stereotypes. English does not have many German loan words - at least not many in common use - but those it does have are a rather more mixed bunch than such stereotypic lists might imply. The following, all relatively current in contemporary English, are German loan words or loan translations pre-1900:

- angst, aspirin, bum (= "lazy person"), chic (< French ?? German schick), Christmas tree, dachshund, death-wish, delicatessen, dollar, doodle, drill (= "cotton cloth"), dumb (= "stupid"), ecology, eiderdown, enzyme, ersatz, fahrenheit, fife (< Pfeife), folk-song,
hamburger, hamster, handbook, heroin, hinterland, hock (= "Rhine wine"), hoodlum, kaput, kindergarten, kohlrabi, lager (= "beer"), larch, noodle, plunder, poltergeist, poodle, rucksack, spanner, swindler, waltz, Yiddish, yodel, zeitgeist.

In addition, German has been a major language of scholarship in the nineteenth and twentieth centuries, particularly in chemistry and other natural sciences, but also in the humanities and social sciences - including linguistics - hence borrowings such as

- biology, chromosome, gauss, gestalt, gneiss, Hertz (Hz), leitmotif/v, marxism, quartz, umlaut, weltanschauung.

SEARCHING BY DATA-BASE AND CORPUS

Major innovations in dictionaries in the last few years provide new ways of studying such loans. Some dictionaries are based on large computer-readable corpora, and contain information on word frequency. Others are available in machine-readable form with search software. For example, the OED on CD-ROM is a relational data-base, which can be searched in many ways, often within seconds. It is possible to search for all words borrowed from a given language, all words first recorded in a given time period, or the combination of the two: say, all German loans first recorded in English after 1900. Clearly it would take years to find such words in the printed OED.

A simple measure of which words are current in contemporary English is whether they occur in the Cobuild (1995) dictionary, since this contains only words common enough in a 200 million word corpus of contemporary English to be included in a dictionary for advanced learners. "Common enough" is defined by frequency in the corpus, and Cobuild codes words in five frequency bands: 700, 1,200, 1,500, 3,200 and 8,100 words each. These words make up 95 per cent of running text in the corpus. The dictionary provides no information on how many head-words outside this total of 14,700 are listed: I estimate a further 15,000. The words listed above, from angst to zeitgeist, are in Cobuild (1995), although mainly not within the 14,700 most frequent words. (See ET 19 and 20 for articles on the 1989 OED in its printed and electronic versions, ET 31 on the use of corpora in making dictionaries, and ET 46 for a review of Cobuild and comparable dictionaries.)

If the OED software is instructed to search for all loans from German after 1900, it finds about 1,250 items. However, this figure is misleadingly high for different reasons. It certainly finds words which are current in contemporary English, including

- abseil, allergy, bakelite, diktat, festschrift, flak, kitsch, nazi, realpolitik, schizophrenia, snorkel, spritzer, stratosphere.

If we are generous about what we include, a search finds up to 75 words which are widely known to native speakers. But this includes many proper names, such as: Alzheimer's [disease], Bauhaus, Dobermann, Geiger [counter], Jugendstil, Rottweiler. And it also includes words which are certainly widely known, but where the etymology is only indirectly German, (e.g. quiche < French < Alsace dialect Küchen < German Kuchen).
A few words name specifically German cultural phenomena, and stereotypes at that, though perhaps \textit{dirndl}, \textit{lederhosen} and \textit{loden} are the only three which are fairly current (and not common enough to be in Cobuild 1995). Jucker (1996) discusses how loan words can provide both positive and negative local colour. One of his examples is the use, in a British newspaper, of a German \textit{autobahn}, rather than \textit{motorway}, to connote high speed, "so that the reader can be astonished at what strange things happen in foreign countries"! (\textit{Autobahn} is in Cobuild 1995.)

If the \textit{OED} software is similarly instructed to search for all German loans from 1800 to 1900, it finds over 2,400 items, but again relatively few of these are current in contemporary English. The following are in the \textit{OED} but not in Cobuild:

- dummkopf, edelweiss, liebfraumilch, ouija (< oui + ja!), riesling, schadenfreude, weltanschauung, weltschmerz.

The \textit{OED} on CD-ROM is a very powerful research tool, which makes possible entirely new investigations of aspects of vocabulary. However, incautious use of the software may produce quantitative results which look superficially plausible - but it may turn out that the computer was not acting very intelligently, because the pattern-matching in the software is rather limited. I have therefore also used other corpora and specialist dictionaries (Algeo 1991, Williams 1976/1983), including the essential historical dictionary by Pfeffer and Cannon (1994), which lists loans not only alphabetically, with etymologies, but also chronologically, and by semantic areas.

\textbf{OTHER C20th GERMAN LOANS}

The figure of 1,250 loans since 1900 found by the \textit{OED} software is a crude statistic, since it does not distinguish between intuitively very different cases.

\textbf{HISTORICALLY MOTIVATED WORDS}

From the Nazi-set above, Cobuild (1995) gives only \textit{blitzkrieg} and \textit{Nazi(sm)}. The German origin of such loans is certainly recognized by English speakers: indeed, that is the whole point of using them. In that sense, they are not English words at all, but German words which can be used in English to refer almost exclusively to a particular historical period. For example, in contemporary German, \textit{Luftwaffe} simply means "airforce". But in English it means "the airforce in Nazi Germany", and is on the hazy boundary between proper names and common nouns. If we are fairly generous about this category, and include other political and military terms (some, I suspect, common only amongst fans of British war films), there are about 25 items. Others are

- gauleiter, herrenvolk, lebensraum, messerschmidt, panzer, stalag, stuka.

In their earliest uses, \textit{blitz}, \textit{blitzkrieg} and \textit{flak} had only a military sense, but they now have general meanings, as in these attested examples from Cobuild (1995):

- there is to be a blitz on incorrect grammar
- a blitzkrieg of media hype
• attracted more than their fair share of flak from the press.

NAMES

Dictionaries are often uncertain how to deal with proper names for people and places, titles of works of art, and so on. The OED software finds about 80 such items, only a few of which are widely known to educated speakers, and their status as words "in" English is perhaps dubious. They include

• Hitler, Humboldtian, Leibniz, Nietzschean, Rheingold (the opera and express train!), Schumannesque.

TECHNICAL TERMS

Any such categorizations will have unclear boundaries, and numbers are tendencies at most. However, the largest category by far (ca. 750 out of 1,250) is technical terms, unknown to most native English speakers. From a random sample of 125 of the 1,250 loans post-1900 found by the OED software, only 7 were in Cobuild. Even if we assume that twice that many are probably widely known by native speakers, this still means that only about 10 per cent of these 1,250 loans are fairly current in contemporary English.

The largest sub-categories (over 30 per cent together) are mineralogy and chemistry, and many other words come from biology, geology, botany, medicine, physics and mathematics. Pfeffer and Cannon (1994) estimate over a longer time period that 50 per cent of German loans are specialist, technical or scientific words. Examples from chemistry, where German language publications were until recently influential internationally, include

• adduct, biotin, dehydrase, emulsoid, heterophile, indigoid, lactol, mutase, perbunan, polyaddition, pterin, sabinene, stigmasterol, uridine, zwitterion.

I gave a random sample of 30 such words to a friendly native English-speaking professor of biochemistry: he knew only half of them.

In addition, many of these words were coined in German from Greek and Latin elements, and are indistinguishable from other Greco-Latin technical terms in English. Indeed, in such cases the notion of etymology - as origin in a given language - is rather dubious. Such words have been formed according to the conventions of professional academic groups, which are largely independent of national or language groups.

LOAN TRANSLATIONS

This category includes

• anti-body << Antikörper, pecking order < Hackliste, (nowadays < Hackordnung), power politics < Machtpolitik, rainforest < Regenwald, space-time < Raumzeit, superman < Übermensch.
The OED notes other cases as merely parallel in German and English: *breakthrough* (*Durchbruch*), *delouse* (*entlausen*), *hookworm* (*Hakenwurm*). If we allow both sets, the software finds about 25 items in this category.

**PANEL 1 - German for chemists**

The influence of German in chemistry is rapidly decreasing. In the 1970s, the University of Nottingham designed a reading course on German for Chemists (Butler et al. 1973), and several British universities still offer such a course. But the need is declining. A leading German chemistry journal, *Angewandte Chemie* (Applied Chemistry), publishes a German version and then an English version, with the same contents, a month later. The German weekly *Die Zeit* reported on 19 July 1996 that two other leading German chemistry journals have decided, from 1997, to cease publication altogether in German and to accept articles only in English.

**INDIRECT LOANS**

The OED software does not distinguish between words loaned recently from German, and words with a more complex history. For example, some words passed from much earlier forms of German into Yiddish - though there may still be comparable words in modern German - and usually via American into British English. There are over 30 examples, such as *bagel*, *glitzy*, *nosh*, *schlep* and *schmal(t)z*. See Rosten (1971).

**GERMAN WORDS**

The software also identifies around 200 words which rather puzzle me: namely, words which just are German! They must have been attested in English texts, or they would not occur in the OED, but it seems highly unlikely that they have any general currency. They include: *hochgeboren* (*high born*, of noble birth), *trockenbeerenlauslese* (*wine from selected over-ripe single grapes*), and *putzfrau* (*cleaning lady, charlady*). Some examples here certainly occur in specialist academic discourse, sometimes because no exact translation is available, and where the referent is a specifically German cultural or political phenomenon. So there is an overlap with the category of specialist terms. Literary examples include: *Künstlerroman* and *Bildungsroman*.

Of course, there can be no definitive list of the words which are "in" English. The vocabulary of a language is an open set, with words coming and going all the time. This is discussed in a famous passage in the original introduction to the OED: the vocabulary of English is "not a fixed quantity circumscribed by definite limits", but rather a "nebulous mass [with a] clear and unmistakeable nucleus [which] shades off on all sides ... to a marginal film that seems to end nowhere".

The vocabulary also shades off into obsolescent and archaic words. Etymological dictionaries typically contain information on the first sighting of a word (usually in a written text), and
loans referring to world political events can often be precisely dated. But we cannot see words die: many of them just crawl away quietly, and hide in large dictionaries, which provide them with last resting places labelled "obsolete" or "archaic". And although words may have a sudden onset of use, they often die out gradually, which is why the language of older speakers may sound old-fashioned. This decreasing frequency is a probabilistic phenomenon, which can only be recorded in large corpora.

FALSE HITS - AND MISSES!

Finally, the software simply makes mistakes. It can identify different, unstandardized ways of referring to German etymology ("G.", "Ger." and "German"), but does not recognize that the word "German" or the abbreviation "G." may occur in the etymological entry for other reasons. It makes about 45 clear errors, such as: Jerry ("probably alteration of German"); and jeep (from "the initials G.P., general purpose")!

The software also certainly misses items, though there is no way of knowing how many - short of reading some 20,000 pages of text (or checking every item against Pfeffer and Cannon 1994). I happen to have noticed the following omissions. Blitz is missed, since the etymology occurs only under blitzkrieg. Diesel is missed, since the etymological information is given, but not in a separate etymology section. Muesli is missed, since the software seems not to recognize the labels "Swiss G." and "Swiss-Ger.". a menu allows a search for "German Swiss", but this doesn't find muesli either. Glitsch (?< glitschen) is given as "etymology unknown". VW is given, but not Volkswagen.

PANEL 2 - Words which didn't make it

The OED gives many German loans which had only a short life in English and didn't gain any general currency. The following are some of my favourites:

- beyondman < Übermensch, cf superman
- double-ganger < Doppelgänger
- merwoman < Meerfrau, "an older or wedded mermaid"
- inwandering < Einwanderung, "action of wandering or straying into some place"
- outwandered < ausgewandert, "emigrated"

FREQUENCY IN THE LEXICON

All in all, if we exclude indirect loans, proper names and the like - and consider separately the large number of highly specialized technical words - we are left with a rather small number of German loan words current in contemporary English.

Pfeffer and Cannon (1994) collected data from the OED plus "all major dictionaries in English", and found many more German loans than previously supposed: over 6,000 from 1500 to the present day. Even this surprisingly large number is only one or two per cent of the English vocabulary as recorded in the OED, which contains over 320,000 head words and
616,500 word forms, including derivatives, phrases, and so on (Algeo 1990). However, Pfeffer and Cannon point out that only some 10 per cent of these loans are in everyday use.

As we have seen, this is because many are scientific and technical words. Therefore, in general corpora, the text frequency of German loans is much lower again. As with the vocabulary as a whole, a few words occur frequently, but most words occur rarely. I searched for all the loans listed above (mainly C20th, and some C19th), including their derivatives (e.g. ecology, ecological, ecologist, etc), in a corpus of 2.7 million words, which contained samples from over a thousand different texts, spoken and written, fiction and non-fiction of many genres, all post-1900, mostly post-1960, and including over 500,000 words from British and American newspapers. Only 15 loans occurred over 10 times:

- antibody, blitz, biology, diesel, dollar, ecology, enzyme, handbook, hinterland, lager, marxism/t, nazi(sm), space-time, waltz, Yiddish.

Most occurred fewer than 5 times: e.g. angst, eiderdown, rucksack, snorkel. Many did not occur at all: e.g. ersatz, kindergarten, poltergeist, zeitgeist.

Even if we include words such as biology (and all its derivatives, frequency 140) and dollar (frequency 80), the combined frequency of occurrence was only about 650. This is fewer than one occurrence every 4,000 words, or about 0.025 per cent of running text.

The text frequency of words depends, of course, on the content of the texts in the corpus. For example, there were 12 occurrences of the loan translation space-time, all from a single text on astrophysics. And there were 38 occurrences of Nazi (nearly 6 per cent of the total of 650), many from texts on German history, though the word is distributed throughout the corpus. However, its frequency does not seem to be merely an artefact of my data, since Cobuild (1995) records Nazi in its third frequency band: amongst the 3,400 most frequent words in English. Even Nazism gets into the top 15,000.

CONCORDANCE DATA

Large corpora can also be used to study word use in more detail. Take the word angst. In everyday German, Angst means "fear, worry, be frightened of ...". There is an OED citation from 1849, but the first main citations are in relation to Freud 1922 and Heidegger 1941. However, in more recent everyday English, it has acquired strong evaluative connotations. Cobuild (1995) notes that it is "used mainly in journalism". In addition, it often occurs in semi-fixed phrases, about 10 per cent in angst-ridden and a further 10 per cent in teen(age) angst, and is often used when other people's behaviour is being evaluated in ironic or highly critical ways. Attested examples (from the 200-million word Cobuild corpus) include:

- piffling, sniffling pseudo-angst
- nasal whining or chestbeating angst
- the usual teen-angst, punkthrash, grungefest boogaloo
- tawdry angst-ridden verses [recollect] his inglorious past
- a tasty mix of angst and vinegary pop thrills
- genuine tortured souls or mere designer angst merchants?
- what lurks at the bottom of the Angst Barrel, be glad we're here to scrape it for you
• self-adjusting cruise control could take much of the Angst out of the Autobahn.

*Angst* is something typically suffered by adolescent, middle-class, introspective, sensitive, suburban souls. The concordance lines in Panel 3 give a few examples of common collocations.

---

**Panel 3 - Concordance for ANGST: selection only**

- She has re-created herself as angst-ridden 15-year-old Lettie Chubb in her Flash Liquid. - June 30: Teen angst evening with The Shangri-Las. - Their fusion gawky energy, teen angst, rap urgency, heavy-duty rhythms, ain repetition emblematic of teen angst. Like Understanding. You don't understand their teenage intensity, hiding their teenage angst. - We're the teenage angst fiery West midlands guitar band, a
- Thrashing out all that teenage angst has meant she's come to know her own autobiographical line, - Teenage angst has paid off well, now I'm bored and perfect with Rollins' teenage angst lyrics. He plays these high-pitched songs the new kid in town - Teenage angst, love, fights, summer camps - they loop of somebody else's teenage angst. They move the goal posts altogether.
- You're looking for a fine line in angst-ridden indie pop with future potentialism of El Penitente or the angst-ridden introspection of Herodiade?
- So begins the dating game, the angst-ridden weeks of plotting and political intrigue. Clive - We're the teenage angst fiery West midlands guitar band, a
- The latest addition to teenage angst. 'I do think about spots, but line up is some truly awful teenage-Angst poetry ('I must write/I must live vampire film with teenage biker angst and Kathryn Bigelow's lovely and e
THE OED AND CULTURAL KEYWORDS

Archbishop Trench (1851), one of the proposers of the OED, saw language as "fossil history". And Raymond Williams (1976), in his famous Vocabulary of Culture and Society, uses the same geological metaphor. Using data from the first printed version of the OED, he investigated the history of over 100 "keywords" in British culture: nodes around which ideological battles are fought. He uses a more modern (Marxist) idiom than Trench, but his basic idea is essentially the same: that words embody facts of history, and that they can be analysed diachronically to reveal unconscious assumptions of their community of users.

Williams shows, for example, that culture has nowadays such a variety of meanings because of semantic extensions in the past, leaving layers of meanings deposited over time. Early meanings, concerning "tending crops or animals" (agriculture), were extended to human development (early 1500s), and to the abstract general process (late 1700s), when the word acquired social class connotations. As a result, we now have different uses, such as sugar-beet culture, germ culture, and culture in the sense of "music, literature and the arts".

Williams also shows that many words in English were influenced by cognate words in French and German. Words are borrowed back and forwards between languages, and although a word-form may not have been borrowed from German, many meanings have been greatly influenced by German-language philosophy, psychology and sociology. In many of his entries, Williams notes the influence of major thinkers such as Freud, Hegel, Herder, Marx and Weber. Cases where Williams notes such German-English relations include:

- aesthetic, alienation, anthropology, bourgeois, capitalism, class, community, culture, dialectic, ecology, ethnic, existential, folk, formalist, genius, history, humanity, idealism, ideology, imperialism, individual, industry, materialism, pragmatic, psychology, romantic, sociology, theory, unconscious.

The word-form unconscious is, of course, not a German loan, but a central part of its contemporary meaning comes from the popularization of Freud from the 1920s onwards. The OED on CD-ROM can also be searched for all quotes from a given author: the software finds half-a-dozen quotes from Freud.

CONCLUSION

Data-bases and large corpora offer vast new resources for studying vocabulary, but care is required in interpreting data which can be collected with such ease. As Jucker (1994: 154) warns, the OED on CD-ROM can be a "dangerous research tool". And one important type of influence cannot be identified at all by the computational methods I have used. The kinds of semantic development and cultural influence studied by Williams (1976) are invisible to a request to the OED software to retrieve "German loans since 1900". Though once one knows what (or who) to look for, it is certainly possible to retrieve information in other ways.

The impact of German on modern everyday English is small, though larger and more varied than often supposed, and the influence is much larger in academic areas. All of this perhaps
does something to balance the stereotyped *blitzkrieg-lederhosen-kitsch* view of German influence on English.

PANEL 4 - German words used in linguistics

The intellectual influence of German-language work is very clear in linguistics: via work, in the 1700s and 1800s by Herder, Humboldt, the Brothers Grimm and others, in the late 1800s by the Junggrammatiker, and by Max Müller, Hermann Paul and Hugo Schuchardt (on creoles), and in the 1900s by, for example, Karl Bühler, Franz Boas and Leonard Bloomfield (a Germanist, of Austrian descent, and greatly influenced in his early work by the German psychologist Wilhelm Wundt). Henry Sweet (1845-1912) spent a year studying at Heidelberg University, emphasized all his life the importance of German linguistic scholarship, and published an EFL book for German learners: *Elementarbuch des gesprochenen Englisch*.

Pfeffer and Cannon (1994) list 101 borrowings in linguistics. The following loans, names and loan translations are all fairly current in English-language linguistics.

ablaut
Ausbausprache = "developed language"
formant
Grimm's Law
High German < Hochdeutsch
isogloss
neo-grammarians < Junggrammatiker
loan translation < Lehnübersetzung
loan word < Lehnwort!
Mischsprache = "mixed language", hybrid language
Plattdeutsch = "Low German"
Rhenish fan < Rheinischer Fächer
schwa < Hebrew sheva; "In German books spelt schwa" (*OED*).
sound shift < Lautverschiebung
Sprachbund = "linguistic area"
Sprachgefühl = "linguistic intuition"
Stammbaum = "family tree"
umlaut
Ursprache = "original language", protolanguage
Urtext = "original text", earliest version of a text
Verner's Law
Wernicke's aphasia
Yiddish
NOTE

A longer version of this article appeared in W. Mackiewicz & D. Wolff, eds (1997) British Studies in Germany. Wissenschaftlicher Verlag Trier. I am grateful to Kirsten Günther for help in extracting data from the OED, and to colleagues at Cobuild in Birmingham for access to The Bank of English corpus.

REFERENCES

DICTIONARIES


OTHER WORK


