Originally published in: International Journal of Lexicography
vol. 33 (2020) no. 2, pp. 150-167.
DOI: *https://doi.org/10.1093/ijl/ecz033*

POSTPRINT

Annette Klosa-Kückelhaus
Sascha Wolfer
Leibniz Institute for the German Language
[klosa|wolfer]@ids-mannheim.de

Considerations on the Acceptance of German Neologisms from the 1990s

Abstract: Are borrowed neologisms accepted more slowly into the German language than German words resulting from the application of word formation rules? This study addresses this question by focusing on two possible indicators for the acceptance of neologisms: a) frequency development of 239 German neologisms from the 1990s (loanwords as well as new words resulting from the application of word formation rules) in the German reference corpus DeReKo and b) frequency development in the use of pragmatic markers ('flags', namely quotation marks and phrases such as sogenannt 'so-called') with these words. In the second part of the article, a psycholinguistic approach to evaluating the (psychological) status of different neologisms and non-words in an experimentally controlled study and plans to carry out interviews in a field test to collect speakers' opinions on the acceptance of the analysed neologisms are outlined. Finally, implications for the lexicographic treatment of both types of neologisms are discussed.

Keywords: neologisms, German, dictionary of neologisms, corpus study, flagging

1. Introduction

Neologisms which are included into dictionaries have presumably been accepted by the language community. There are some (neologism) dictionaries that record neologisms at their 'moment of birth' (cf. Lemnitzer 2010: 67), that is before they are fully lexicalized, but nevertheless accepted parts of the lexicon (e.g. the German neologism dictionary *Die Wortwarte*). There are others – and all general language dictionaries as well as the German *Neologismenwörterbuch* (see Klosa and Lüngen 2018, Steffens 2017, al-Wadi 2017) at the Leibniz Institute for the German Language belong into this category – that record

neologisms only after their lexicalization, that is retrospectively. In both types of neologism dictionaries as well as in general language dictionaries, new borrowings as well as new words resulting from word formation rules are included as headwords, if the dictionary in question follows a descriptive approach. However, from a puristic point of view and therefore within a prescriptive approach, a dictionary may limit or exclude borrowings (cf. Zgusta 1971: 183-184). From a prescriptive point of view, one might argue that borrowings are not accepted as easily and/or quickly as new words resulting from the application of word formation rules,[1] thus justifying their exclusion from a dictionary.

In the remainder of the article, we will label the latter group of neologisms – namely those that are the result of applying German word formation rules to German language material – *German neologisms*. This is merely a label for the sake of brevity. We also view the other group, namely the borrowings, as part of the German language. In our paper, we will test the hypothesis that borrowings are accepted more slowly, also raising the question of how to operationalize 'acceptance'. Finally, we will look at the implications of how to treat neologisms of both types in (descriptive) neologism and general language dictionaries.

We will present the first steps of our research to answer the question of which linguistic markers are used as indicators for the acceptance of neologisms, for example, whether increasing frequency is effectively related to acceptance. First, we will focus on approximately 133 German word formation products (compound nouns such as *Blitzeis* 'black ice' or *Schaltfläche* 'button on a web page') and 102 borrowed nouns (e.g. *Booklet*, from Engl. *booklet* or *Ciabatta*, 'Italian white bread', from Italian *ciabatta*) from the beginning of the 1990s. These lexemes are fully described in the *Neologismenwörterbuch*. We will discuss their overall frequency development in the German reference corpus DeReKo and examine in detail the use of pragmatic markers, which are also called *flags* (cf. Poplack and Sankoff 1988: 1178), with these neologisms.

In addition to this corpus-based study, we will present our plans to carry out interviews in a field test to collect speakers' opinions on the acceptance of the analysed neologisms. In this way, we will be able to see whether our findings based on corpus data converge with the ratings and evaluations from German speakers. We will also outline a psycholinguistic approach to evaluating the (psychological) status of different neologisms and non-words in an experimentally controlled study. Finally, we will discuss whether questions of the acceptance of neologisms by speakers of a language (here: German) should have an impact on the inclusion of these words in or exclusion of these words from different dictionary types, especially a general monolingual dictionary in contrast to a dictionary of neologisms.

2. Acceptance of neologisms

New words occur in German all the time, but not all of these are neologisms. Many new compounds or derivations are used only once in (spoken or written) texts; thus, they are not lexicalized, but nonce words. Neologisms, on the other hand, are lexical units or meanings which emerge in a communication community in a specific period of time of language development, which diffuse, are generally accepted as language norm, and which the majority of speakers perceive as new for some time (cf. Herberg et al. 2004: XII). According to this definition, neologisms are entrenched parts of the German lexicon, they are 'felt to be a full member' (Haspelmath 2009: 43) of German. The problem with this definition is, of course, that we need criteria to measure entrenchment, that is whether a new word has been lexicalized and is indeed 'generally accepted as language norm'. As possible indicators

for the acceptance of a neologism in German, criteria such as the following can be taken into account:

a) increasing overall frequency of a neologism,
b) distribution of the neologism in many different text types,
c) use of the neologism in many different discourses.

Other indicators, which tell us how far the process of lexicalization of new words has developed, are (cf. Lemnitzer 2010: 69):

d) Pragmatic criteria: A neologism is no longer written in quotation marks, hedge words (e.g. *sogenannt* 'so-called') are not used any more, distancing phrases (e.g. *wie man heute sagt* 'as we say today') are abandoned.
e) Grammatical criteria: The gender of nouns is invariable, a full conjugational paradigm for verbs is developed.
f) Word formation criteria: A neologism is used as first and second component in an increasing number of compound nouns, a borrowed neologism is combined with indigene lexemes in word formation products.

Taking these criteria into account, we should be able to measure significant differences in our corpus, for example in increasing overall frequency or in how quickly the above-mentioned markers for neologisms are abandoned, if the acceptance of borrowings is lower than the acceptance of German neologisms. In the following, we will focus on increasing overall frequency and the use of quotation marks and some distancing phrases as indicators of the degree of lexicalization and acceptance of neologisms. However, we will neither look into the extension of the distribution of new words in different text types and discourses, nor discuss grammatical criteria and the productivity of neologisms in word formation, but reserve these for future research. In addition, we are not currently pursuing further suggestions for indicators of lexical entrenchment made by Chesley and Baayen (2010: 1344), who suggest examining a borrowing's dispersion ('the number of text chunks a word occurs in if a text is divided into several subparts'), length ('we hypothesize that the length of the borrowing will be inversely related to its degree of entrenchment') and sense pattern ('monosemy vs. polysemy'), as well as the cultural context in which it is used ('whether or not the borrowing in a particular context refers to the culture of the language of the borrowing') to learn about possible differences in the acceptance (or entrenchment into the German lexicon) of borrowings vs. German neologisms.

When we assume that there is a relationship between increasing overall frequency of a (borrowed) neologism and its acceptance, we follow a more general assumption of diffusion: loanwords 'typically increase in frequency and diffuse across speakers. This assumption in turn appears to stem from the inferences that (1) a nonce word can be equated with the first stage of lexical innovation, (2) more frequently occurring foreign words represent a later stage in this development [...].' (Poplack and Dion 2012: 285). We also consider the increase in frequency to be an indicator of the lexicalization of a neologism formed (by composition or derivation) in German. For both types, but particularly for borrowings, we presuppose that their use is generally accepted by the German language community (even if some puristic attitudes, especially against Anglicisms, are articulated): 'If successful establishment of loanwords depends on their frequency of use [...], then it must make some difference whether or not it is deemed okay to use foreign words at all' (Backus 2014: 25).

We also assume that there is an 'inverse relationship between flagging and frequency' (Grant-Russel and Beaudet 1999: 26), meaning that with increasing frequency of a new lexical item the use of flags with this word decreases. This in turn suggests that 'the more a borrowing is perceived as standard usage in the language community addressed by the discourse, the more it is unmarked' (Grant-Russel and Beaudet 1999: 26). This assumption also applies to the use of quotation marks with neologisms: Palmer and Harris (1990: 83) postulate for the use of English loans in French that when the use of quotation marks with these borrowings is given up, this 'indicates either a high degree of acceptance of these terms or that the writers of these documents are unconscious of these as English terms'. We will see whether this is also true for German neologisms in our data. Looking more specifically at the expression *sogenannt* 'so-called' (the most common distancing phrase used with neologisms in German), it is certainly true that 'In its default use, *sogenannt* informs the addressee about the status of the head nominal as a conventionalized term used in a certain speech community' (Härtl 2018: 141). We will look at the development in the use of *sogenannt* 'so-called' and similar expressions, both with borrowed neologisms in German and with German neologisms, to see whether there are any differences between these neologism types.

3. Frequency development of German neologisms in DeReKo

All analyses have been carried out using the statistical software environment *R* (R Core Team 2019). All visualizations have been created using the package *ggplot2* (Wickham 2016).

3.1 *First test*

Our first step was to test the frequency development of 35 noun neologisms (15 borrowings, e.g. *Anchor, Cybersex, Zapping*, and 20 German compound nouns, e.g. *Erlebnisgesellschaft* 'fun- and pleasure-seeking society', *Fahrstuhlmannschaft* 'sports team moving up and down between leagues', *Schaltfläche* 'button on a web page') first attested in the German language between 1991 and 1993 and coming from a similar frequency class (frequency class 20 in the German Reference Corpus, DeReKo).[2] In a lemmatized search, we found 76,641 results for the whole set, which consisted of 46,074 results (60.1%) for borrowed and 30,567 results (39.9%) for German word formation products. We tried to find test words with a roughly similar frequency development. However, this turned out to be very difficult when selecting words by overall frequency class alone. Of course, words from the same frequency class can have very different frequency developments through time. This is also the case in our sample. Thus, we discarded this criterion for our further studies. The accumulated development of the simple relative frequencies of our first test set as illustrated in Figure 1 shows no clearly distinguishable pattern in the frequency development for borrowed neologisms vs. German neologisms.

3.2 *Expanded database*

The second step was to extract more neologisms from the beginning of the 1990s and add them to the database, leading to a set of 134 compound nouns and 105 borrowed nouns. We used a lemma-based search and considered potential alternative spellings listed in the *Neologismenwörterbuch*. After investigating the results, we had to exclude four words from the final dataset for word-specific reasons: *Doppelpass* ('double ID') as a neologism refers to the ID documents of people with dual citizenship. However, *Doppelpass* also

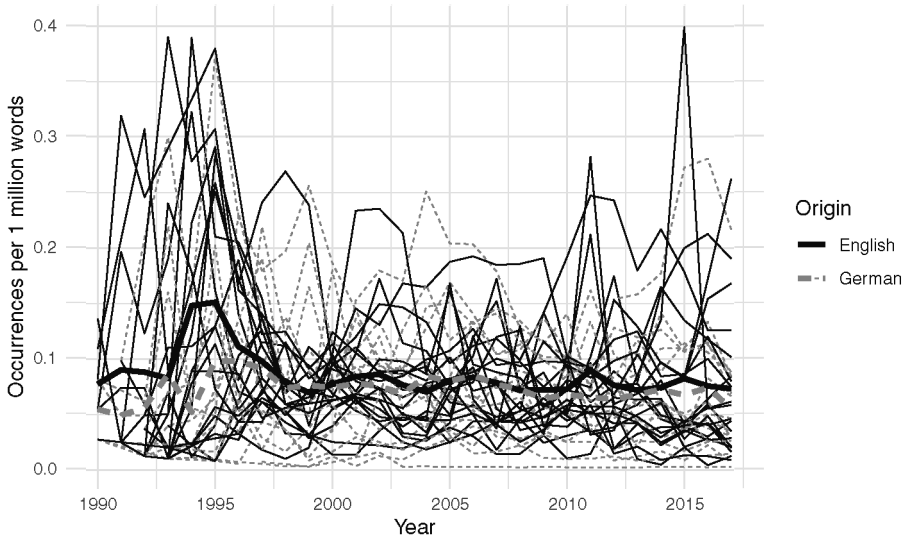

Figure 1. Frequency developments for words in the first test set of German neologisms from the 1990s (solid black: loanwords from English, dotted grey: German compound nouns). Thick lines represent yearly averages.

refers to the much more frequently used 'one-two pass', e.g. in football, which is not a neologism at all. *IT* as an abbreviation for *Informationstechnologie* ('information technology') had to be excluded because there are numerous occurrences of *IT* as an abbreviation for 'Italy', e.g. after the names of athletes. Another word we had to exclude was *Loser* (which has the same meaning as in English) because it also frequently occurs as a family name in the corpus, the same is true for *Gate*. The remaining words (again, we searched for lemmas) appeared 5,369,834 times in the corpus with 3,159,141 occurrences of borrowed forms (58.8%) and 2,210,693 occurrences of German word formation products (41.2%) after 1979. We chose this year as a cut-off point to see whether we could find results that lie before the period described by the *Neologismenwörterbuch*.

Figure 2 illustrates the frequency development of the 20 most frequent nouns in this dataset (both borrowed and formed in German) since the beginning of the 1990s. While some of these show a steady increase overall (e.g. *Event*, *Hüpfburg* 'bouncy castle', *Location*), others show larger deflections in frequency development (e.g. *Beachvolleyball*, *Mobiltelefon* 'mobile phone'), and some show extreme peaks (e.g. *Techno*, *Mausklick* 'mouse click'), but no overall frequency decrease, thus supporting our initial findings from the smaller dataset: there is no clearly distinguishable pattern that would separate borrowings from German neologisms.

But when we look at our expanded database as a whole, we have to revise this first impression. Figure 3 illustrates the frequency development of all neologisms investigated through time with aggregated values for the two categories and associated standard errors. Some words are evidently attested before 1990, but because they show a clear frequency increase at the beginning of the 1990s, they are classified as neologisms of this decade in the *Neologismenwörterbuch* from which our search terms were taken. In contrast to our observation based solely on the test words, here we see a clear difference between the overall

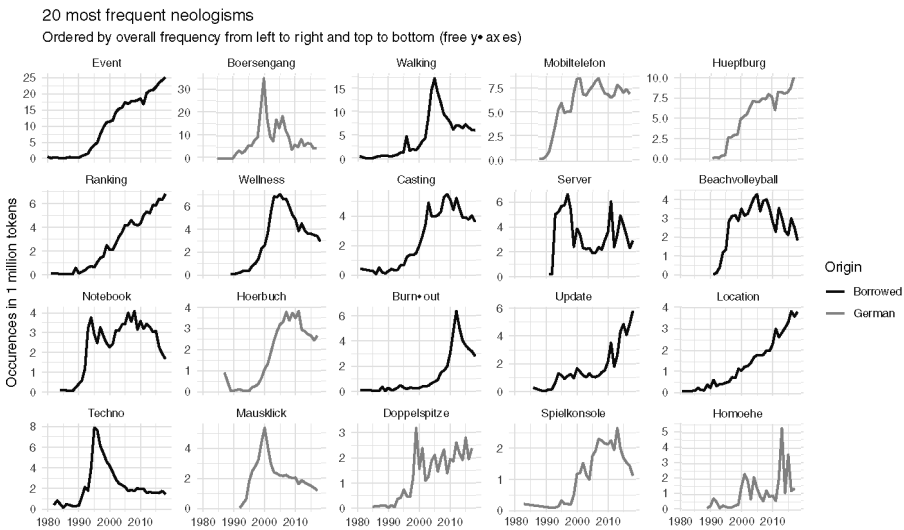

Figure 2. Frequency development of the 20 most frequent German neologisms from the 1990s in the expanded database (black: loanwords from English, grey: German compound nouns). Please note that the *y*-axes vary freely between panels. This means that, for example, *Event* (with the *y*-axis going up to 25 occurrences in 1 million tokens) is much more frequent overall than, for example, *Notebook* (*y*-axis up to 4 occurrences in 1 million tokens).

frequency development of borrowed neologisms vs. German neologisms. While the German compound nouns investigated in our corpus have on average been on the decline since they emerged, the borrowed nouns are on average not only found more frequently in our corpus, but also seem to have stayed at a certain level of use. If we accept the assumption that higher frequency indicates a later stage of lexicalization of a neologism, we could draw the conclusion that the borrowed nouns in our dataset seem to be fully lexicalized and well accepted in German, even when only taking a simple measure such as relative frequencies into account. As a side note: *Event* (which has the same meaning as in English) and *Börsengang* ('initial public offering') are very frequent in our database and it might well be the case that the very frequent lemmas dominate the pattern we see in Figure 3. However, when these two lemmas are excluded from the analysis, the differences between borrowed neologisms and those of German origin do not change considerably (see Figure 3, right panel).

So, given the words in our dataset and frequency development as an indicator, we have to assume that borrowings are not accepted less than German neologisms. However, frequency development can be influenced by a wide variety of factors. To focus more closely on acceptance, we investigated the use of linguistic markers (or flagging) within our dataset next. As described in Section 2, flagging can also be seen as an indicator of how accepted a term is by the speech community.

4. Use of linguistic markers with German neologisms in DeReKo

In the following, we no longer distinguish between the first test and the expanded database. This allows us to focus more on the questions at hand and not on the differences between

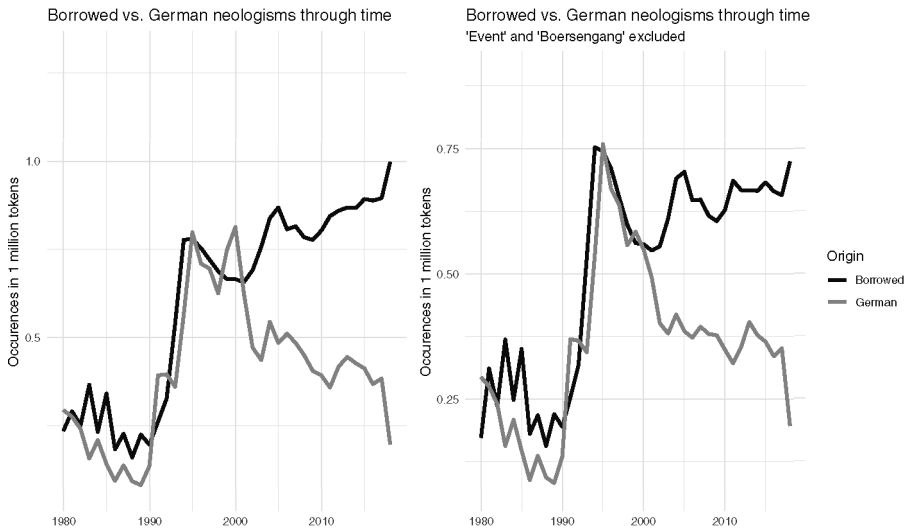

Figure 3. Aggregated frequency development of German neologisms from the 1990s (expanded dataset; black: loanwords from English, grey: German compound nouns; shaded areas indicate standard errors; on the left: all words; on the right: *Event* and *Börsengang* 'initial public offering' excluded)

the pilot study and the full set of neologisms. We also excluded all results from the database before 1990 to focus on the time period covered by the *Neologismenwörterbuch*. This leaves us with a dataset of 4,578,163 occurrences (55.7% borrowed, 44.3% German).

We considered seven linguistic markers (shown in Table 1). We searched for these markers within a window of ten words around the relevant keyword. Whether we searched before and/or after the keyword is also indicated in Table 1. For some markers, it was necessary to restrict the search to only one side of the neologism because some markers make sense on one side but not the other. For example, it would not make sense to search for all forms of *das heißt* ('that means', 'i.e.') **before** the keyword because the keyword would then be part of the explanation and not the explicandum itself. But only the latter case is an indication that the writer felt the need to explain the word to the reader – a case of the linguistic marking we are looking for.

We also tested some other potential markers (e.g. *Anglizismus* ['Anglicism'] before or after the keyword or parentheses following the keyword) that we had to exclude at a later stage because they either simply did not occur in the vicinity of our keywords (in the case of *Anglizismus*) or were prone to 'false alarms', i.e. they were not a clear sign of linguistic marking. (Parentheses, for example, occur frequently in the vicinity of our keywords, but only in very few cases contain an explanation or comment concerning the keyword.) Due to the large amount of data we processed, we did not have the opportunity to manually go through all hits but had to draw small samples to check whether a marker was valid. So, we had to concentrate on markers that are either already known in the literature (e.g. Lemnitzer 2010: 69) or are otherwise clear indicators of linguistic marking.

Table 1 also contains the number of occurrences for each marker. It is very clear that quotation marks are the flagging device that is used most often. As a side note: we considered German single and double quotes as well as French double quotes (« ») because the

Table 1. Investigated linguistic markers ('flags') for German neologisms of the 1990s

Marker (German)	Marker (English)	Occurrences
quotes around the keyword	quotes around the keyword	172,943
sogenannt, so genannt, sog. before keyword	'so-called' (with abbreviation)	37,438
zum Beispiel, z. B. after keyword	'for example', 'e.g.'	8,993
genannt after keyword	'called'	6,072
englisch before or after keyword	'English'	5,127
neudeutsch before or after keyword	'new German'	1,561
das heißt, das bedeutet, d. h. after keyword	'that means', 'i.e.'	1,004
Number of marked keyword occurrences (one or more of the above)		226,551

latter are not consistently standardized to German quotes in the corpus. Since our results are so dominated by quotation marks, we do not analyse the different markers separately but treat an occurrence of a keyword as marked as soon as one or more of the above markers appear in its vicinity. That is why the last row of Table 1 is not the sum of all markers. For example, an occurrence could be marked by *sogenannt* and could be quoted at the same time. (This combination occurred 3,707 times in our database.)

First, we will contrast the number of marked occurrences for borrowed and German neologisms. Next, we will focus on the relationship between frequency and linguistic marking. Finally, we will investigate how borrowed and German neologisms from between 1990 and 2017 are marked.

4.1 *Flagging of borrowed vs. German neologisms*

Of the 2,551,412 occurrences of the borrowed neologisms, 160,206 (6.28%) were accompanied by at least one linguistic marker. German neologisms were marked in 3.27% of all cases (66,345 of 2,026,751 occurrences). Markings vary widely for the different keywords. To assess this variation, we calculated the ratio of marked occurrences for each keyword by dividing the number of marked occurrences by the number of all occurrences. We use this dataset for the following analysis. Table 2 shows the ten neologisms of each group with the highest ratios of marked occurrences.

There are several ways to test whether the two groups substantially differ in their flagging status. We provide two ways of testing: a logistic regression based on the individual occurrences of all words ($n = 4{,}578{,}163$) and a comparison of the group means, based on an aggregated keyword list ($n = 235$). In the logistic regression, the probability of being marked is predicted by membership of one of the two groups (borrowed vs. German). Such an analysis takes all individual occurrences into account and measures how much more (or less) probable it is that an occurrence is marked if the keyword belongs to the German group. The result of this analysis shows that it is significantly less probable that a keyword is marked if it belongs to the German group ($\beta = -0.683$, $SE = 0.0047$, $z = -144.8$, $p < .0001$). For the comparison of the group means, we calculated a one-sided permutation test which is a non-parametric and, generally speaking, more robust version of the t-test. A one-sided test is justified in this case because the hypothesis laid out at the beginning of the paper states that borrowed neologisms are less accepted than German neologisms (and should therefore be marked more often). The permutation test with 50,000 Monte Carlo replications also suggests a meaningful difference between the two groups of neologisms

Table 2. The ten neologisms of each group with the highest ratios of marked occurrences in the corpus. The rightmost column provides a translation of the German neologisms (if available).

Borrowed	Ratio	German	Ratio	Translation
Cocooning	44.3%	*Dezemberfieber*	56.6%	December fever
Mc-Job	42.8%	*Buschzulage*	52.8%	a bonus payed to somebody who is willing to work in unattractive regions
Updating	42.0%	*Crashkid*[3]	49.7%	joyrider
Late Show	40.6%	*Ereignisfernsehen*	49.5%	event TV
Generation X	36.2%	*Knopflochchirurgie*	44.1%	keyhole surgery
Get-together	35.0%	*Leihbeamter*	41.6%	loan/borrowed official
Lean Production	34.3%	*Schlüssellochchirurgie*	38.7%	keyhole surgery
Edutainment	31.5%	*Tigerland*	32.5%	tiger state (Asian tiger)
Novel Food	30.2%	*Täterakte*	31.9%	offender file
Dreamteam	27.3%	*Wossi*[4]	30.6%	<no translation>

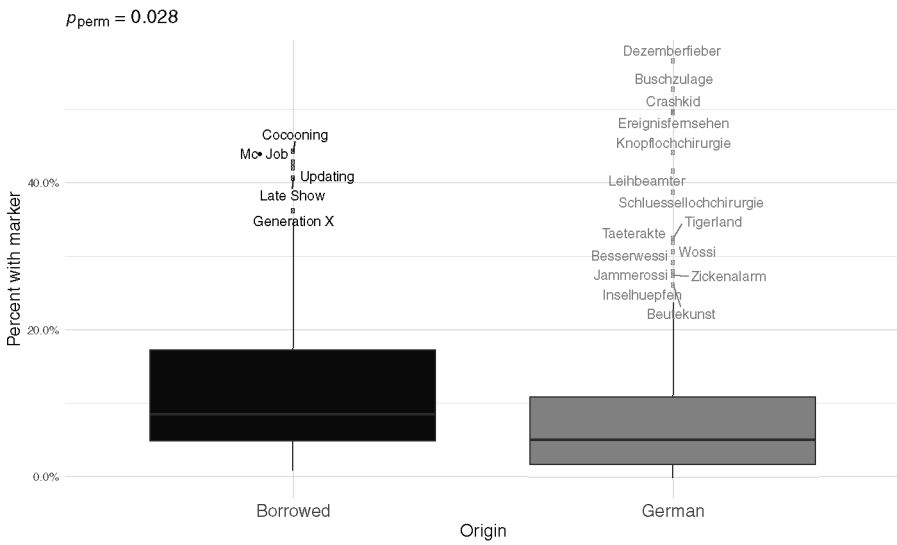

Figure 4. Boxplot of the ratio of marked occurrences for borrowed and German neologisms. Outliers are indicated as dots in their specific locations with associated labels. 50% of the keywords lie within the relevant box; medians are indicated by the horizontal line within the boxes.

($p = .028$) with borrowed neologisms being marked more often. Both testing methods indicate a significant difference between the two groups. Hence, we can conclude that borrowed neologisms are indeed marked more often. Figure 4 visualizes this difference by means of a boxplot; outliers are indicated in the figure. Group medians are 8.57% for the 'Borrowed' group and 5.07% for the 'German' group.

In summary, we can conclude that the difference between the two groups is not large (mean difference of 3.01 p.p. for the individual occurrences and 2.88 p.p. for the

aggregated dataset) but statistically significant. In the next section, we will investigate whether marking of the neologisms is related to frequency of occurrences in the corpus.

4.2 *Relationship between linguistic markers and frequency*

The most straightforward way to test whether marking and frequency of use are related is to calculate a correlation coefficient. We are using Spearman's rank correlation coefficient ρ because (unlike Pearson's r) it does not make a linearity assumption for the relationship between the two variables. As a statistical test for significance, we are again using a permutation-based p-value (based on 10,000 Monte Carlo replications). When we correlate the frequency of occurrence in the corpus with the ratio of marked occurrences for all 235 keywords, the test signals a medium negative correlation of $\rho = -.367$ with a permutation-based $p < .0001$. This means that keywords that occur more often in the corpus are clearly less likely to be marked. This also holds for both groups separately, i.e. this effect is not driven by one group alone (borrowed: $\rho = -.405$, $p < .0001$; German: $\rho = -.446$, $p < .0001$).

Correlations might sometimes be 'triggered' by outliers in the data. In our case, there are some keywords that are much more frequent than others. For example, as already shown in Section 3.2, *Event* and *Börsengang* occur very often (375,954 and 294,759 times respectively). But even when we restrict the database to neologisms with an overall frequency of less than 20,000 (which leaves 193 neologisms), we still observe a highly significant negative correlation ($\rho = -.282$, $p = .0001$) between frequency and marking ratio.

4.3 *Linguistic markers through time*

Figure 5 shows the development of linguistic marking through time for the 20 neologisms that appear most often in the corpus. These are the same ones as in Figure 2, only the variable shown on the y-axis has now changed and the time period (x-axis) begins in 1990. For some words (e.g. *Location* in the lower left-hand corner of the plot), there are only a few occurrences in earlier years which can lead to some 'jumpy' behaviour of the line graph because just a few marked occurrences lead to big changes in the marking ratio. Not all neologisms behave like the very prototypical example *Burn-out*: this neologism starts out with a high marking ratio which decreases steadily over time. The same can be said of *Event*, *Walking*, *Wellness*, *Location* and *Techno*. There are other words, such as *Ranking*, *Notebook* and *Beachvolleyball*, which show a rapid decline in markedness. So, there is not one kind of development that could be generalized to all investigated neologisms. If we want to draw more general conclusions, we have to aggregate the values over all patterns to see whether a general pattern emerges that distinguishes borrowed from German neologisms.

Figure 6 visualizes such an aggregation by calculating the group average per year. We can clearly see that this average value of marking ratio is consistently higher for borrowed neologisms than for German ones. In Section 4.1, we already established that this is not a huge difference but it is statistically meaningful. Another thing we can see in Figure 6 is that there is a steady decline in markings during time. In 2017, the mean marking ratio for borrowed neologisms is only 5.25%; the mean value for German is 4.31%. There is also a hint of an interaction effect in Figure 6, i.e. borrowed neologisms decline faster than German ones with regard to linguistic marking.

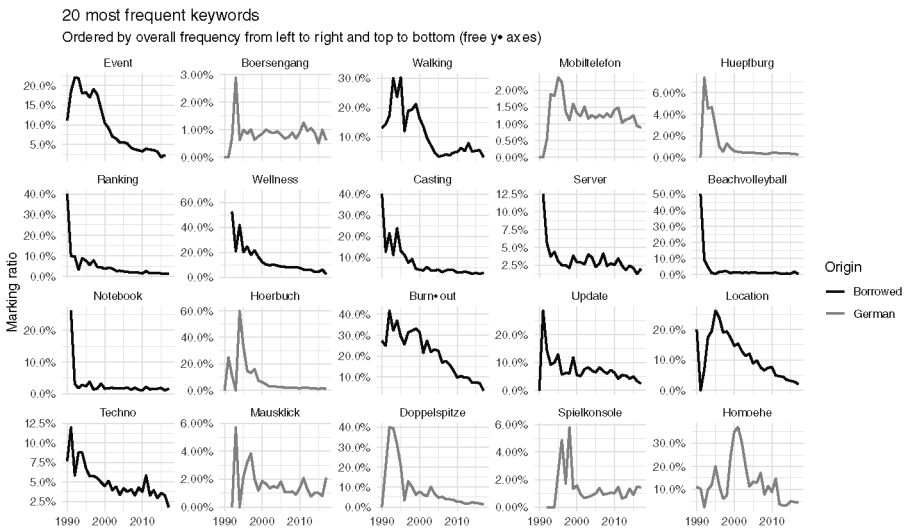

Figure 5. Development of linguistic marking between 1990 and 2017 for the 20 most frequent neologisms in our database. Please note that the *y*-axes vary between panels.

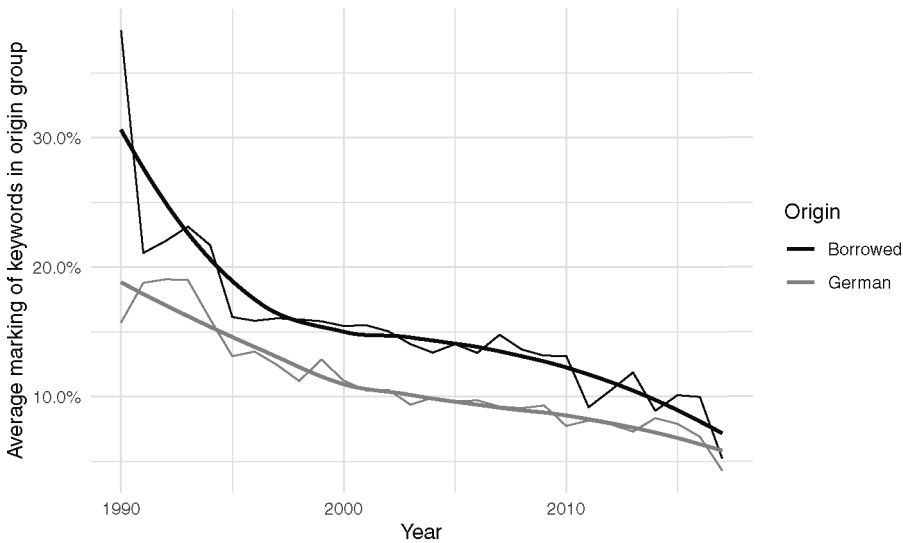

Figure 6. Aggregated development of linguistic marking between 1990 and 2017 for all neologisms in our database. Aggregation is done over all keywords in each group. The thin lines represent raw data. The thick lines are a local smoothing (LOESS) line taking 75% of the data into consideration at any point.

To corroborate this observation, we calculated a linear mixed-effects model using the lme4 package (Bates et al. 2015) which also takes differences between the keywords into account by including them as so-called random intercepts. The results suggest a negative relationship between year and flagging frequency ($\beta = -0.00561$, $t = -19.4$), overall fewer

markings of German neologisms ($\beta = -2.61$, $t = -3.34$) and an interaction between year and origin ($\beta = 0.00129$, $t = 3.29$), indicating that marking of borrowed neologisms is indeed abandoned faster than for German neologisms. None of the coefficient's confidence intervals include 0. Hence, we have to assume that these effects are significant. The interaction can also be expressed by the following statements. Borrowed neologisms start out at a higher level of markedness. (The average marking of borrowed neologisms between 1990 and 1994 is 24.1% vs. 17.9% for German neologisms.) Consequently, there is more 'room' for them to be unmarked than for the German neologisms. At the end of the time period, both types of neologisms are much closer together in terms of marking (9.26% borrowed vs. 6.97% for German between the years 2013 to 2017.). This 'difference in differences' is captured by the interaction.

4.4 *Conclusions*

In this section we found clear differences between borrowed and German neologisms when it comes to being accompanied by linguistic markers. German neologisms are marked less often, when the whole timeline from 1990 to 2017 is taken into account. Also, neologisms that appear more often in the corpus tend to be marked less often. An interesting pattern emerges when we look at the time course of flagging neologisms. Although it is true that borrowed neologisms are marked more often overall, this effect seems to be restricted to the beginning of the period we investigated. After roughly 30 years, the difference is barely perceptible. This suggests that after some time, both borrowed and German neologisms are 'accepted' in the language to an equal extent.

5. Plans for further studies

With the results from our corpus study in mind, we plan to carry out interviews in a field test to collect speakers' opinions on the acceptance of a sample of the neologisms we analysed. In this way, we will be able to see whether our findings based on corpus data converge with the ratings and evaluations from German speakers. Alternatively, we could follow Soares da Silva (2014) who tested 'whether speakers' knowledge of the origins of words corresponds to actual language behavior' (2014: 127) and set up a survey in which 'both cognitive and behavioral factors of language attitudes' were combined. Participants were presented with a set of alternative expressions for one concept and were asked to choose the one they 'would *usually/sometimes/never* use' in standard Portuguese. We could adapt this idea to test whether borrowed neologisms or German neologisms are more likely to be used in standard German and correlate the results with our corpus findings. In the following, we will not elaborate these ideas any further, but instead outline a psycholinguistic approach to evaluate the (psychological) status of different neologisms and non-words in an experimentally controlled study.

We intend to set up an experiment with the software MouseTracker which records and analyses mouse movements travelling towards potential on-screen responses (cf. Zybatow and Weskott 2018). This method will allow us to gain insights into potential 'insecurities' of participants concerning two alternative answers. Table 3 illustrates the set-up of the experiment combining two factors: origin (German compound noun vs. noun borrowed from English) and status as a neologism (neologism from the 1990s vs. neologism from the 2010s vs. non-word). For each test word, a participant will be able to click on either *yes* or *no* on the screen to answer the question 'Is the word you see a *real/good/accepted* word?'.

Table 3. Potential set-up of MouseTracker study on the acceptance of German neologisms

		Factor A: Neologism status		
		90s neologism	**2010s neologism**	**non-word**
Factor B: Origin	**German**	*Trennkost* 'food-combining diet'	*Schwarmstadt* 'city attracting mainly young people'	*Durstmagnet* 'thirst magnet'
	English	*Tanktop*	*Shapewear*	*Doorapp*

The software MouseTracker will record the movement of the mouse towards the potential answers. Our hypotheses are: we should receive more negative answers for the neologisms from the 2010s, because they are not yet fully accepted; we should receive even more negative answers for non-words as these cannot be lexicalized at all, but are in fact nonce words; we would also assume that we would see higher deviations from the optimal mouse path towards one of the answers for newer neologisms as these are not as frequent and established as the older ones from the 1990s; finally, we would expect to see lower deviations from the optimal mouse path (towards *no*) for non-words as their status should be not questioned at all. The tricky part, of course, will be choosing good examples, creating convincing non-words and finding neutral wording for the question above to minimize the risk of undesirable influences on our participants.

6. Neologisms in dictionaries

In this section, we will consider briefly whether questions of the acceptance of neologisms by speakers of a language (here: German) should have an impact on the inclusion of these words in or exclusion of these words from different dictionary types. As we have seen from our corpus data, both borrowed neologisms and German neologisms are well accepted into the German language, especially after the new words have been present in the language for some time. Thus, both types seem to be candidates for inclusion in monolingual and bilingual as well as general and specialized dictionaries.

Accepting a neologism as a headword in a dictionary mainly depends, of course, on the inclusion criteria established for the dictionary in question. For example, a general monolingual dictionary will apply different criteria from a specialized dictionary of neologisms. While in the general monolingual dictionary only highly frequent neologisms from general language may be explained, in a specialized dictionary such as the *Neologismenwörterbuch* less frequent ones or some from special vocabulary may also be included, if frequency is one criterion for the inclusion/exclusion of headwords. If, on the other hand, puristic views determine the selection of headwords, borrowed neologisms might be excluded from a dictionary. As the acceptance of borrowed neologisms in German according to our results seems not to be lower than the acceptance of new German words resulting from the application of word formation rules, for descriptive dictionaries of contemporary German we suggest treating both neologism types equally.

Notes

1. In the call for papers for the international conference 'New Words and Linguistic Purism' (Innsbruck, 25-26 October, 2018) this question was put as follows: 'To what

extent are borrowings resisted more strongly than new words resulting from the application of word formation rules?', presupposing that the acceptance of borrowings is lower than the acceptance of new words resulting from the application of word formation rules.

2. Frequency class 20 means that the most frequent word *der* is approximately 2^{20} times more frequent than all words belonging to this frequency class.
3. It could be argued that *Crashkid* is not formed from German material and should thus be counted as a borrowed neologism. However, it is a pseudo-Anglicism (the English equivalent is 'joyrider') and we therefore did not count it as borrowed and assigned it to the German group.
4. Explaining all the neologisms is outside the scope of this article. However, we would like to give a brief explanation of *Wossi* because there is no English translation at all. *Wossi* is a combination of *Wessi* (someone from western Germany, the Federal Republic of Germany) and *Ossi* (someone from eastern Germany, the former German Democratic Republic) and refers to a person who moved from the West to the East after German reunification in 1990.

References

A. Dictionaries

Neologismenwörterbuch. Accessed on 24 April 2019. http://www.owid.de/wb/neo/start.thml.
Die Wortwarte. Accessed on 26 September 2019. http://wortwarte.de.

B. Other literature

Backus, A. 2014. 'A Usage-Based Approach to Borrowability.' In Zenner E. and G. Kristiansen (eds), *New Perspectives on Lexical Borrowing. Onomasiological, Methodological and Phraseological Innovations*. Boston/Berlin: De Gruyter Mouton, 19-39.

Bates, D., M. Maechler, B. Bolker and St. Walker. 2015. 'Fitting Linear Mixed-Effects Models Using lme4.' *Journal of Statistical Software* 67.1: 1-48. doi:10.18637/jss.v067.i01.

Chesley, P. and R. H. Baayen. 2010. 'Predicting New Words from New Words: Lexical Borrowings in French.' *Linguistics* 48.4: 1343-1374.

Grant-Russel, P. and C. Beaudet. 1999. 'Lexical Borrowings from French in Written Quebec English.' *University of Pennsylvania Working Papers in Linguistics* 6.2: 17-33.

Härtl, H. 2018. 'Name-informing and Distancing *sogenannt* 'so-called': Name Mentioning and the Lexicon-pragmatics Interface.' *Zeitschrift für Sprachwissenschaft* 37.2: 139-169.

Haspelmath, M. 2009. 'Lexical Borrowing: Concepts and Issues.' In Haspelmath M. and U. Tadmor (eds), *Loanwords in the World's Languages. A Comparative Handbook*. Berlin: De Gruyter Mouton, 35-54.

Herberg, D., M. Kinne and D. Steffens. 2004. *Neuer Wortschatz. Neologismen der 90er Jahre im Deutschen*. In Collaboration with E. Tellenbach and D. al-Wadi. Berlin/New York: de Gruyter.

Klosa, A. and H. Lüngen. 2018. 'New German Words. Detection and Description.' In Čibej, J., V. Gorjanc, I. Kosem and S. Krek (eds), *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts*, Ljubljana: Ljubljana University Press, Faculty of Arts, 559-569.

Lemnitzer, L. 2010. 'Neologismenlexikographie und das Internet.' *Lexicographica* 26: 65-78.

Palmer, J. D. and D. Harris. 1990. 'Prestige Differential and Language Change.' *Bulletin of the CAAL* 12.1: 77-86.

Poplack. S. and N. Dion. 2012. 'Myths and Facts about Loanword Development.' *Language Variation and Change* 24: 279-315.

Poplack, S. and D. Sankoff. 1988. 'Code-switching.' In Ammon, U. N. Dittmar and K. J. Mattheier (eds), *Sociolinguistics: an International Handbook of the Science of Language and Society, Vol.* 2, Berlin/Boston: De Gruyter Mouton, 1174-1180.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Soares da Silva, A. 2014. 'Measuring and Comparing the Use and Success of Loanwords in Portugal and Brazil: a Corpus-based and Concept-based Sociolectometrical Approach.' In Zenner, E. and G. Kristiansen (eds), *New Perspectives on Lexical Borrowing. Onomasiological, Methodological and Phraseological Innovations.* Boston/Berlin: De Gruyter Mouton, 101-141.

Steffens, D. 2017. 'Vom Print- zum Onlinewörterbuch – Zur Erfassung, Beschreibung und Präsentation von Neologismen am IDS.' In Dąbrowska-Burkhardt, J., L. M. Eichinger and U. Itakura (eds), *Deutsch: lokal – regional – global.* Tübingen: Narr, 281-294.

al-Wadi, D. 2017. 'Begegnungen mit neuen Wörtern: Zu lexikografischen Praktiken im *Neologismenwörterbuch* des IDS.' In Dąbrowska-Burkhardt, J., L. M. Eichinger and U. Itakura (eds), *Deutsch: lokal – regional – global.* Tübingen: Narr, 173-186.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer.

Zgusta, L. 1971. *Manual of Lexicography.* The Hague/Paris: Mouton.

Zybatow, T. and T. Weskott. 2018. Das Doppelperfekt: Theorie und Empirie. *Zeitschrift für Sprachwissenschaft* 31.1, 83-124.

Appendix 1: First test set of neologisms of the 1990s

Compound noun	Borrowed noun
Boxenluder	Anchor
Erlebnisgesellschaft	Cocooning
Farhstuhlmannschaft	Cybersex
Gelbrotsperre	Edutainment
Induktionsherd	Emoticon
Inselhüpfen	Fundraiser
Konsumraum	Girlgroup
Kuschelrock	Homeshopping
Lebensabschnittspartner	Mousepad
Lotuseffekt	Onliner
Mauszeiger	Outplacement
Personenmine	Splatter
Quotenkönig	Streetwear
Rentenlücke	Upgrading
Schaltfläche	Zapping
Sozialbetrüger	
Sparmobil	
Theraband	
Wohlfühlgewicht	
Zickenalarm	

Appendix 2: Expanded dataset of neologisms of the 1990s

Compound noun	Borrowed noun
Ankerwährung	Fengshui
Armutsfalle	Karaoke
Atomkoffer	Anchorman
Autoteilen	Anchorwoman
Bauchgefühl	Animalprint
Besserwessi	Assessmentcenter
Beutekunst	Barcode
Billigjob	Baseballcap
Blitzeis	Beachvolleyball
Börsengang	Blockbuster
Bungeespringen	Bodypainting
Bürgergeld	Bodypiercing
Buschzulage	Booklet
Couchkartoffel	Bungeejumping
Crashkid	Burn-out
Dateiformat	Callanetics
Datenautobahn	Canyoning
Datenhandschuh	Carsharer
Datenhighway	Carsharing
Dezemberfieber	Cashcow
Doppelpass	Casting
Doppelspitze	Couchpotato
Dreiliterauto	Cybernaut
Druckraum	Cybersex
Einheitswährung	Daily Soap
Einkaufsmall	Dreamteam
Ereignisfernsehen	Edutainment
Erlebnisgesellschaft	Electronic Banking
Eurogeld	Electronic Cash
Eurowährung	Event
Eventmarketing	Flipchart
Fixerraum	Fundraiser
Fixerstube	Fundraising
Flachbildschirm	Gameshow
Formatradio	Gate
Freisprechanlage	Generation X
Freisprecheinrichtung	Get-together
Gauckbehörde	Girlgroup
Gelbsperre	Give-away
Genfood	Global Player
Genmais	Global Village
Gentomate	Homebanking
Gerechtigkeitslücke	Hypertext
Gesundheitsraum	Icon
Handtelefon	Infotainer

(continued)

(continued)

Compound noun	Borrowed noun
Häppchenjournalismus	Infotainment
Homoehe	IT
Hörbuch	Late Show
Hüpfburg	Late-Night-Show
Jammerossi	Lean Management
Jobticket	Lean Production
Kerneuropa	Location
Kinderkanal	Loser
Knopflochchirurgie	Mall
Konsensgesellschaft	Management-Buy-out
Konsensgespräch	Mc-Job
Konsolenspiel	Multiplex
Krankenversichertenkarte	Nightshow
Kreativdirektor	Notebook
Kunstevent	Novel Food
Kuschelrock	Outing
Leihbeamter	Outsourcing
Magnetgleiter	Pager
Mauerschütze	PDA
Mausklick	Political Correctness
Mauszeiger	Pumpgun
Minusrunde	Rafting
Mobbingberatung	Ranking
Mobbingtelefon	Rave
Mobiltelefon	Reality-TV
Multiplexkino	Realityshow
Nachwendezeit	Riverrafting
Opferakte	Sequel
Ostalgie	Server
Osterweiterung	Showview
Ostidentität	Sitcom
Outdoorjacke	Smartcard
Outdoorsport	Splatter
Politsprech	Splatterfilm
Pop-up-Buch	Splattermovie
Reformstau	Streetball
Rinderwahnsinn	Talkradio
Schlüssellochchirurgie	Techno
Schwangerenkonfliktberatung	Teleworking
Schwangerschaftskonfliktberatung	Touchscreen
Schwarzkonto	Trackball
Seitenaufprallschutz	Turn-around
Sektdusche	Update
Semesterticket	Updating
Shoppingmall	Upgrade
Solidaritätszuschlag	Upgrading

(continued)

(continued)

Compound noun	Borrowed noun
Solidarzuschlag	Virtual Reality
Solizuschlag	Walking
Soundkarte	Wellness
Sozialbetrug	Windbreaker
Sozialmissbrauch	Womanizer
Sparauto	Zapping
Sparmobil	
Spartenkanal	
Spartensender	
Spielkonsole	
Steckzigarette	
Strandvolleyball	
Superzahl	
Täterakte	
Technomusik	
Telebanking	
Telefonbanking	
Tigerland	
Tigerstaat	
Vereinigungskriminalität	
Versichertenkarte	
Wegfahrsperre	
Weichei	
Werbeinsel	
Werbepause	
Wohlfühlgewicht	
Wossi	
Zeitkonto	