Yasmin Crombez*, Anne-Sophie Ghyselen,
Esme Winter-Froemel and Eline Zenner

The sociopragmatic parameters steering the reported selection of Anglicisms or their Dutch alternatives

https://doi.org/10.1515/ling-2020-0237
Received October 28, 2020; accepted April 19, 2021; published online May 10, 2022

Abstract: Researchers studying language variation and change induced by contact with English initially focused on the linguistic integration of English source language (SL) material in the morphophonological structure of the receptor language (RL). Shifting the attention toward the sociopragmatic localization of English lexical material, researchers now foreground both the social identity work conducted with English material and the pragmatic functions of English lexemes and heritage alternatives. So far, however, most of these studies rely on production data, paying less attention to language users' perception and evaluation of English lexemes. Therefore, integrating insights from language production and evaluation, this study asks the questions (i) which lexical preference speakers express when asked to choose between an English and a heritage form; and (ii) which sociopragmatic parameters help to explain this preference. To answer these questions, a forced-choice experiment was conducted in which over 1,500 Dutch-speaking participants were asked to select their preferred expression when provided with a loanword and a heritage alternative. Each participant was offered 12 target trials and 10 filler trials. The target trials, drawn from a corpus-based study, include a balanced set of loans from three semantic fields with various frequencies in usage data. Pragmatic variation was included by randomly presenting the trials in speech contexts of communicative immediacy and communicative distance. Mixed-effects logistic regression analyses

***Corresponding author: Yasmin Crombez**, Center for Linguistics, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium, E-mail: yasmin.crombez@vub.be. https://orcid.org/0000-0001-5742-6190
Anne-Sophie Ghyselen, Linguistics Department, Universiteit Gent, Gent, Belgium, E-mail: annesophie.ghyselen@ugent.be. https://orcid.org/0000-0001-8242-7419
Esme Winter-Froemel, Neuphilologisches Institut/Romanistik, Universität Würzburg, Würzburg, Germany, E-mail: esme.winter-froemel@uni-wuerzburg.de. https://orcid.org/0000-0003-3437-8864
Eline Zenner, Quantitative Lexicology and Variational Linguistics, Katholieke Universiteit Leuven, Brussel, Belgium, E-mail: eline.zenner@kuleuven.be. https://orcid.org/0000-0002-8114-5425

show a significant impact of community-based usage statistics, the age and self-reported attitude of the participant, as well as the speech context in which the concepts were offered.

Keywords: contact linguistics; Dutch; English loanwords; lexical choice; pragmatics

1 Introduction

Language variation and change induced by contact with English has become a prominent research topic within linguistics over the past decades. This holds for both intense contact situations, where English is a dominant language because of immigration or colonization (cf. Wolf and Polzenhagen 2009), and for weak contact situations, where English does not have official status, and contact with the language occurs more indirectly, predominantly through the media (cf. Androutsopoulos 2012). In the latter case, scholars have mainly focused on lexical borrowing, as that is the most common and apparent result of language contact (Thomason 2001: 69).

Anglicism research, which foregrounds lexical borrowing from English, has historically typically been conducted from a systemic and structural perspective. Most attention so far has been paid to the position of the loanwords in the structure of the recipient language (see Zenner and Kristiansen 2014: 1), resulting in descriptions of the different morphophonological and orthographic adaptation processes (Núñez Nogueroles 2017; Saugera 2017) or the establishment of a borrowability hierarchy comparing different word classes (Czech-Rogoyska and Krawiec 2018). More recently, however, two paradigm shifts in Anglicism research can be remarked: from a systemic to a more sociopragmatic and usage-based perspective (see, among others, Andersen 2014; Ilić 2017; Onysko and Winter-Froemel 2011; Onysko et al. 2013; Peterson 2017; Peterson and Vaattovaara 2014; Winter-Froemel 2011: 295–319, 2017), and, as a result, from a semasiological to an onomasiological point of view (Soares da Silva 2014; Zenner and Kristiansen 2014; Zenner et al. 2012).

The first shift "implies a reorientation of its locus from the borrowed lexemes per se, to how the use of borrowed items is constrained by cultural, social or cognitive factors" (Andersen et al. 2017: 71). This means that researchers now look at lexical borrowing as an instrument of local identity construction, speaker evaluation communication, and as manifesting sociocultural norms and values, or in short: as socially meaningful (Andersen 2014; Babel 2016; Peterson and Beers Fägersten 2018; Zenner et al. 2019). The second shift refers to taking the concept expressed by the Anglicism, rather than the lexical borrowing itself, as the point of

departure, suggesting that attention is paid to the naming instead of the meaning, and as such to the selection made by language users between the loanword and alternative lexicalizations (see for instance Winter-Froemel [2008] who gives an overview of criteria for judging between alternative lexical strategies in contact situations, and Serigos [2017] and Winter-Froemel [2018] who show that Anglicisms are often semantically more specific than their Spanish and Italian alternatives using a concept-based approach).

In adopting these sociopragmatic perspectives on lexical variation induced by contact with English, most researchers have concentrated on the production of Anglicisms in various contexts and genres, while fairly exclusively relying on corpus-based methods (Balteiro 2011; Burmasova 2010; Yang 1990). Experimental studies, targeting language users' perception and (social) evaluation of Anglicisms, are as yet far more scarce, with studies combining information from corpus data and experimental data virtually absent altogether (though see Sandøy [2014] as an exception). Nevertheless, a proper sociopragmatic account of lexical borrowing that enhances our understanding of the lexical choices language users make between loanwords and heritage alternatives, is contingent on a combination of insights from language production, perception and (social) evaluation (Backus 2020).

Against this background, the objectives of the present study are twofold. First, we aim to combine both the sociopragmatic and the onomasiological shift by investigating the under-explored issue of lexical choice between loanwords and heritage alternatives. Second, in selecting the appropriate methodology to obtain the first objective, this study targets to complement the strongly corpus-based orientation of Anglicism research by relying on an experimental design, though not without bearing the fruit of existing corpus-based Anglicism studies. Specifically, we investigate the sociopragmatic parameters steering the lexical preference of Belgian Dutch and Netherlandic Dutch speakers for English lexemes or Dutch alternatives through a forced-choice experiment that is designed based on the corpus study of Zenner et al. (2012).

By focusing on English loanwords in the Low Countries, this article presents Anglicism research in a weak contact setting. While most speakers of Dutch claim to be fairly bilingual according to the *Eurobarometer* of 2012 (European Commission 2012), English does not have official status in either country. Consequently, contact with English happens predominantly indirectly via marketing, television, movies, music, and other mass and social media (De Decker and Vandekerckhove 2013; Zenner et al. 2015). The Low Countries are additionally an interesting case study given the pluricentric nature of the Dutch language (Clyne 1992). It is sometimes argued that the Flemish have a different relationship with English than the Dutch because of Flanders' history and

sociolinguistic background (see e.g., Geeraerts and Grondelaers 2000; van der Sijs 1996; Zenner et al. 2012, 2015). Whereas the Netherlands developed a Dutch standard language from the seventeenth century onwards, the social elite and the ruling class in Flanders preferred French as their overt language of prestige. As a consequence, Geeraerts and Grondelaers (2000: 58) argue that "the struggle for recognition of Dutch as the official language in Belgium, often materialized as a competition with the French standard", culminated in the rejection of French interference, and, even possibly foreign interference altogether (Vandenbussche et al. [2005: 46], and see also Cajot [2004]). While this has not yet been corroborated empirically to our knowledge, this study might shed light on the possible impact of the different linguistic histories of the Low Countries on self-reported preference toward Anglicisms.

The paper is structured as follows: first, we provide an overview of existing works on the perception and (social) evaluation of Anglicisms in Section 2. In the third section, we introduce the research questions and hypotheses. Section 4 proceeds to the design of our study, where the materials and stimuli of our forced-choice experiment will be presented and the sample will be described. Next, the results are explained in Section 5, followed by a discussion and conclusion in Section 6.

2 The perception and (social) evaluation of Anglicisms

As mentioned above, lexical borrowing from English has mainly been documented through corpus-based studies, with a clear focus on production (see e.g., Andersen and Graedler [2020] for English borrowings in Norwegian; Balteiro [2011] for sports Anglicisms in Spanish; De Decker and Vandekerckhove [2013] for the use of English in Flemish youth speak; Garcia-Yeste [2013] for English loans in Swedish advertising; Martin [2008] for the presence of English in French advertising; Schaefer [2019] for English in German, and many others). Experimental designs targeting perception and (social) evaluation of Anglicisms are less frequent, but not entirely absent.

Van Meurs et al. (2014), for instance, designed an association task to study to what extent speakers perceive Anglicisms to have different meanings from their Dutch alternatives. To study the social meaning of English swearwords in Finnish, Vaattovaara and Peterson (2019) modified the classic matched guise technique from Lambert et al. (1960)'s Speaker Evaluation Paradigm by including availability and acceptability ratings. The study of Sandøy (2014) forms a notable exception as

well, as it investigates the (social) evaluation of Anglicisms in the Nordic languages by combining newspaper data, interviews, and a matched/verbal guise test.

Whether a speaker has a positive or negative reaction to language variation induced by contact with English depends on various factors, social, pragmatic and linguistic. When investigating the attitudes of Nordic speakers toward English influence, for instance, Thøgersen (2004) discovered that positive attitudes toward Anglicisms correlate with social class. The higher the degree of formal education, the more positive Nordic speakers are oriented toward English influx (Thøgersen 2004: 35).[1] Devos (2018) on the other hand, found age to have a significant impact: Flemish participants between 15 and 18 years old accepted the use of English in newscasts more than participants between 60 and 75 years old. A Finnish study then, conducted by Sánchez and Tuomainen (2014), showed Finnish women to be more appreciative of Anglicisms than Finnish men, conforming to the "gender paradox", which states that women often initiate linguistic change yet are more sensitive to linguistic norms (Labov 2001). Besides typical social predictors, context can play a crucial role in shaping attitudes as well (see Garrett [2010] for a general appraisal). An investigation on the acquisition of loanword prestige – via a matched guise method – discovered that Flemish primary school children prefer a cartoon with English loanwords slightly more than a cartoon with only Dutch words (Zenner et al. 2020). While this does not mean that children find English prestigious overall, it does show that in the specific framework of a cartoon, the use of English words is evaluated positively.

Not only in cartoons, but also in sports (cf. Balteiro 2011), job advertisements and commercials (cf. van Meurs et al. 2007), the copious use of English is often explained by referring to its prestige value. It is argued that the English language is associated with modernity and globalization and as such offers a certain appeal for a product (Cheshire and Moser 1994; Martin 2007; Piller 2001). This proposition appears to be confirmed by the research of Ager (1999), who analyzed the attitudes toward English of French native speakers using the results of the SOFRES poll of 1994.[2] 41% of the respondents described Anglicisms as "modern," 30% as "useful," and 19% as "amusing." A follow-up study by Walsh (2015), however, indicated that, although the participants displayed a generally positive stance toward English lexemes in French, over two-thirds agreed that the use of Anglicisms feels superfluous in product packaging (68%), advertisements (69%) and television/radio (61%). Similar results were found by van Meurs et al. (2017), who studied how English loanwords were perceived by Dutch and Flemish students in

1 Participants from Iceland form an exception to these results (Thøgersen 2004).

2 Société française d'études par sondages [French Society of Survey Studies].

job titles. They concluded that English borrowings were largely reacted positively to, but both nationalities preferred the Dutch job titles over the English alternatives nonetheless. These findings support the more general conclusion that English loanwords are appreciated on the one hand because they seem to carry prestige, especially when there is no native alternative, but are rejected on the other because they feel like a threat to the native language, precisely because there is a native equivalent (Sánchez and Tuomainen 2014).

When looking at the existing work on the perception and (social) evaluation of Anglicisms, three methodological oppositions can be made. First, where some of these existing studies rely on direct attitude measurement such as surveys (e.g., Thøgersen 2004), interviews (e.g., Sánchez and Tuomainen 2014), and forced-choice tests (e.g., Walsh 2015), others rely on indirect techniques, typically the matched-guise technique from Lambert et al. (1960)'s Speaker Evaluation Paradigm (e.g., Vaattovaara and Peterson [2019]; Zenner et al. [2020]; or see Sandøy [2014] who relies on the verbal guise technique). These indirect procedures are often preferred over direct methods as they allow an assessment of covert, unconscious attitudes: the participants do not know that their language evaluation is being investigated. While the matched guise technique is indeed a very popular tool for language attitude research, it has been a target of critique as well. Besides the fact that it is hard to find a speaker who can speak multiple varieties naturally, one could also question whether the goal is to keep the informants unaware of one speaker using different language varieties or to keep the informants unaware of evaluating language variation (Pharao and Kristiansen 2019). In that regard, direct acceptability judgments can be rather useful as participants are consciously offering their language evaluations, although one has to keep in mind that these might be steered by a need for being socially or politically correct and that they might not match the actual usage of the participants.

Second, from a usage-based perspective, we can discern between "artificial" and more "naturalistic" research on the perception and (social) evaluation of lexical borrowing. Studies where the participant has to read or listen to a fragment (see e.g., Gerritsen et al. 2000) are generally considered to have a more natural feel because a "real-life" sociopragmatic context is available. This is in opposition to trial experiments, such as a forced-choice task, where stimuli are presented in isolation most of the time, giving the study a more artificial feel (see e.g., Rothe 2014). Nevertheless, as Rosseel et al. (2019) rightfully argue, contextual features need to be included in experimental designs either way, as context plays a crucial role in shaping language evaluations and perceptions as we have discussed above. Hence, this study will attempt to incorporate the context of use into its experimental design.

Third, we find that where some scholars focus on the perception and (social) evaluation of particular Anglicisms, others adopt a rather broad approach aggregating over different Anglicisms. For instance, to study the social meaning of English swear words, Vaattovaara and Peterson (2019) solely focused on the perception of the English loanword *shit* and the Finnish heritage form *paska*. They designed a matched-guise technique with six utterances containing either *shit* or *paska*, all with the same semantic meaning, but spoken in different varieties. Van Meurs et al. (2014) on the other hand, presented Dutch native speakers with 30 English – Dutch word pairs used in job advertisements to investigate whether loanwords receive different meanings than their native equivalents. If we want to gain more insight into the lexical choices of speakers, however, it is necessary to find a golden mean between specificity and aggregation: on the one hand, enough stimuli should be included to allow conclusions transcending the individual lexeme, but on the other hand, stimuli have to be carefully chosen to unveil the factors impacting lexical choice. For this study specifically, this resulted in 12 Anglicisms that are equally spread over different categories according to the different linguistic variables that are under scrutiny (see Section 4.2.1).

Thus, to enhance our understanding of language users' *in situ* selection of loanwords over native alternatives, we propose a forced-choice experiment based on and linked up with corpus data, in which Dutch-speaking respondents are asked to select the most appropriate word from several word pairs consisting of an English loanword and a Dutch alternative, according to a specified context. This way, we can pay attention to both the social and the pragmatic effects steering that choice. The specific research questions and hypotheses underlying the experiment are presented in the next section.

3 Research questions and hypotheses

The main research question underlying this study is: "Which sociopragmatic factors help explain language users' reported preference for English loanwords or their Dutch alternatives?". To tackle this question, we presented a sample of over 1,000 Belgian Dutch and Netherlandic Dutch respondents with several experimental trials in which they were asked to select which of two words they find most appropriate in one of two speech contexts (a quality newspaper article or a conversation with friends): an English loanword or a Dutch alternative. As will be explained in Section 4, the experimental design is set up in such a way that the impact of three sets of sociopragmatic factors on the lexical choice made by the respondents can be determined, viz. (i) the community-based usage patterns for the borrowed and heritage form under scrutiny as established in previous

corpus-based research (see Zenner et al. 2012); (ii) the socio-demographic characteristics of the respondents participating in the experiment; (iii) the speech context offered as part of the experimental trials. As such, our main research question can be subdivided into three subquestions.

RQ1: Do community-based usage statistics for the word pairs under scrutiny help explain the respondents' reported preference for English loanwords or their Dutch alternatives in the experiment?

For the first subquestion, we aim to inquire whether the proportions of use of the English loanwords vis-à-vis the offered Dutch alternatives as established in previous corpus-based research relying on mass media sources help explain the lexical choices made by the respondents in our study. Following the usage-based hypothesis (Goldberg 2006; Tomasello 2005) which states that our mental lexicon is shaped by individual usage events, we can expect the corpus-based proportions to effectively correlate moderately to strongly with the choices respondents make in our forced-choice experiment. Although the link between frequency of use and entrenchment should not be overstated (Schmid 2010), we do anticipate language users to find the lexeme that occurs most frequently to be the most appropriate in use. The association between corpus frequencies and self-reported preference in the experiment might however not be absolute, given that purist reflexes might steer the respondents away from borrowed lexical items, even when these are the most frequent option (see e.g., Schmidt and Diemer 2015; Spitzmüller 2007; Walsh 2014).

RQ2: Do the socio-demographic characteristics of the respondents help explain their reported preference for English loanwords or their Dutch alternatives in the experiment?

Four socio-demographic characteristics are considered in this study, viz. the respondents' age, gender, region, and self-reported attitude toward the influence of the English language on Dutch. As concerns Age (RQ 2.1), we anticipate that younger respondents will be most likely to select the English loanword over the Dutch alternative and that older respondents will be least likely to select the English loanwords over the Dutch alternatives. We base this hypothesis on claims in previous research that the use of English words and phrases is a youth language phenomenon (see De Decker and Vandekerckhove [2013] and Devos [2018] for Dutch; Garley [2010] and Pennycook [2003] on English as the language of hip hop, and also more broadly Piller [2001]) and on research revealing more

purist reflexes in older respondents (e.g., Walsh [2015] for French). As concerns Gender (RQ 2.2), previous research shows conflicting results. Where González Cruz and Rodriguez Medina (2011) do not find any notable gender differences, Zenner et al. (2015) reveal higher use of English by male participants on a Dutch reality TV show (later shown as indexing a strong heterosexual male identity, see Zenner et al. [2015]). Sánchez and Tuomainen (2014) on the other hand show women as more appreciative of Anglicisms than men. This is further supported by Balteiro (2014), who highlights the important role of English lexemes in fashion discourse, which is still typically female oriented. Hence, no clear hypothesis for gender effects can be put forward, but the results from previous studies do point out that more research on the potentially gendered nature of the use of English loanwords is needed. To investigate Region effects (RQ 2.3), we contrast Belgian Dutch and Netherlandic Dutch respondents, expecting a stronger purist anti-English reflex for the Belgian Dutch participants. Although to date no synchronic empirical support has been found to support this hypothesis (despite attempts in, among others, Zenner et al. [2012, 2015]), we anticipate that the standardization history of Dutch in Flanders (the Northern part of Belgium) as described in Section 1, will leave its trails in the language attitudes of the Belgian Dutch respondents (see Cajot [2004] on purist reflexes in Flanders, which in his view materialize especially in complaints about French loanwords in Netherlandic Dutch, rather than in the actual language usage in Flanders). It is, however, also possible that the described puristic reflexes mainly pertain to French, given Flanders' linguistic history, and less so to English. Finally, we include the respondents' Self-reported attitude toward the influence of English on Dutch (RQ 2.4), verifying to what extent a negative attitude toward English leads to a stronger penalization of English loanwords.

RQ3: Do speech contexts of communicative immediacy versus distance help explain the respondents' reported preference for English loanwords or their Dutch alternatives in the experiment?

Previous research has highlighted that the evaluation of particular Anglicisms may vary according to the contexts in which the items are used, that is, speakers may rate the adequacy of an Anglicism compared to its heritage alternative differently depending on contextual features (Zenner et al. 2020). This observation can be linked to the variation linguistics framework developed by Koch and Oesterreicher (2001, 2012), where communicative immediacy and communicative distance are introduced as a basic dimension of linguistic variation. Scenarios of immediacy and distance are characterized by parameters such as the dialogic versus

monologic character of communication, the familiarity versus unfamiliarity of the communication partners, face-to-face interaction versus communication across temporal and spatial distance, private versus public communication, spontaneous communication marked by expressivity and affectivity versus planned communication on a previously fixed subject, etc. Moreover, communicative immediacy is typically realized in oral settings, whereas distance communication is traditionally linked to written realization. A general crosslinguistic observation is that for scenarios of communicative distance, there is a stronger orientation to the standard language. In our view, this also implies a more conservative attitude and an avoidance of Anglicisms, as purism is a very important part of standardization (cf. Langer and Nesse 2012: 612; Van der Sijs 1999: 111).

In addition to that, immediacy and distance are linked to other dimensions of lectal variation, e.g., informal peer group conversations are perceived as belonging to communicative immediacy in a broad sense, whereas formal style, found in newspaper articles, for instance, is linked to communicative distance. To investigate the potential impact of communicative immediacy and distance on the evaluation of loanwords or their Dutch alternatives, we, therefore, operationalize this distinction by presenting the experimental trials (see Section 4) in two different usage contexts: a quality newspaper article or a conversation between friends. Whereas the first context exhibits typical features of communicative distance such as the monological, premeditated and public character of the utterance, etc., the informal conversation with friends represents one of the most typical scenarios of communicative immediacy (private, dialogic, spontaneous, face-to-face communication, etc.). Given the associations of English with modernity discussed above (see Piller 2001) and the general perceptions of communicative immediacy and distance, we anticipate that respondents will be more likely to select the English loanword over the Dutch alternative when asked to imagine a conversation between friends than when asked to imagine a newspaper article.

4 Methodology

To tackle the research questions, we use data from a forced-choice experiment that was embedded in an online survey. In this section, we outline the design of our questionnaire (Section 4.1), the stimuli and materials used (Section 4.2), the procedure of the study (Section 4.3), the participants (Section 4.4), and the analyses (Section 4.5).

4.1 Design

To gather the data for this study, we distributed an online survey, made up of four components. The core of the experiment was a forced-choice task, which prompted respondents to choose one of two synonymous lexemes as most fitting for a given context. A forced-choice test was chosen because it allows us to come as close as empirically possible to capturing respondents' language preferences in "normal, actual usage situations (i.e., linguistic performance in production)" (Arppe and Järvikivi 2007: 151). To distract respondents from the essence of the task, viz. the English – Dutch oppositions, several synonymous word pairs containing only Dutch lexemes as well as established French loanwords, were offered as filler items. Overall, our forced-choice test was set up as a mixed design: each respondent received all word pairs once in a random order, in one of two randomly assigned contexts. By presenting the questions and answers options in a randomized order we hoped to avoid priming. This means that Participant X could be offered Stimulus Y in Speech Context 1 as the first question, while Participant Z could receive Stimulus Y in Speech Context 2 as the third question.

The second part of the survey consisted of two sanity checks meant to verify the core assumptions that guided the stimulus selection: (i) that the word pairs offered in the trials are perceived as synonymous; (ii) that the English loanwords offered in the trials are perceived as foreign items. The third part of the survey openly questioned the respondents about their relationship with English and in the final part, respondents filled out their socio-demographic information.

4.2 Materials

In the following section, the materials used in each of the four parts of the survey are discussed in more depth. First, the details of the forced-choice experiment are described (Section 4.2.1). Second, we illustrate the workings of the sanity checks (Section 4.2.2), after which the questions about the respondents' relationship with English (Section 4.2.3) and their socio-demographic information (Section 4.2.4) are exemplified.

4.2.1 Forced-choice experiment

As mentioned above, the forced-choice experiment contained 22 synonymous word pairs, including 12 target trials and 10 filler trials. A wildcard was introduced

with the option "I don't know (one of) these words" in case the informants did not know the concept or one of the words (see (1)).

Before the participants started the forced-choice task, they were informed on the introduction page of the survey that they would be presented with synonymous word pairs out of which they had to choose the synonym they found most appropriate according to a certain context. To make sure all participants knew what we meant by "synonymous word pairs," a Dutch example of two synonymous words was given, viz. *schoonbroer* versus *zwager* "brother in law," which was accompanied by the specific instruction found in (1). During the forced-choice task, at the top of each page, a short title said: "*kiezen tussen synoniemen*" "choosing between synonyms" under which the instruction and answer options were displayed.

As can be seen from (1), we offered the word pairs in either a context of communicative immediacy, operationalized as "most appropriate in a conversation with friends" or a context of communicative distance operationalized as "most appropriate in a quality newspaper." To ensure maximal understanding and focus, the semantic field to which the concept belonged was emphasized as well.[3] The order of the questions, as well as the answer options, were randomized to avoid priming.

(1) Example of a target trial
Welk woord vindt u het meest geschikt in een gesprek met vrienden [CONTEXT] over sport [SEMANTIC FIELD]? ("Which word do you find most appropriate in a conversation with friends [CONTEXT] about sports [SEMANTIC FIELD]?")
- Keeper [LEXICALIZATION 1]
- Doelman [LEXICALIZATION 2]
- Ik ken (een van) deze woorden niet ("I don't know (one of) these words" [WILDCARD])

The 12 target trials consisted of 12 Anglicisms that occur in Dutch alongside their Dutch alternative. These Anglicisms were chosen based on the corpus study of Zenner et al. (2012), which investigated the success of 149 English person reference nouns in Flemish and Dutch newspapers. As many previous borrowability studies researching the impact of part of speech have already established that nouns are most borrowable, Zenner et al. (2012) decided to focus on the factors explaining variation in borrowability and the success of loanwords within the class of nouns. They opted for person reference nouns specifically as these "have the added advantage that they are similar enough to be compared, yet come from a variety of

3 Note that the information presented between square brackets in Example (2) was not included in the actual survey, but is added here for a better understanding of the set-up of the trials.

different semantic fields" (Zenner et al. 2012: 758). Using a concept-based approach, they defined the success of an English loanword as the relative preference for the Anglicism vis-à-vis existing synonymous expressions (Zenner et al. 2012: 753). To clarify: in the corpus study, the success rate of the English loanword *keeper* was calculated by dividing the number of times *keeper* occurred (32,606) by the number of times the native alternatives *doelman*, *doelverdediger*, *doelwachter* and *kiep* as well as *keeper* occurred (195,738) (= corpus frequency), which equals a success rate of 16.66%. To tackle the issue of polysemy, expressions with multiple meanings which had a low token frequency were manually disambiguated; expressions with a high token frequency were semi-automatically disambiguated using context cues. High frequent concepts for which reliable disambiguation was not possible were discarded from the study (see Zenner et al. 2012: 764).

The selection of the 12 Anglicisms used in this study from the available set of 149 results from a two-tiered procedure. First, we controlled for age of borrowing and relative length of the loanword: all Anglicisms were borrowed before 1989 and had the same word length (in syllables) as the Dutch alternative.[4] We specifically included Anglicisms that were borrowed before 1989, as the results of Zenner et al. (2012) showed that Anglicisms borrowed after 1989 are significantly less successful. As such, by only including loanwords before 1989, we tried to eliminate the possibility of certain Anglicisms being less successful simply because they were borrowed after 1989 (see discussion section for further reflection). If concepts had more than one Dutch alternative listed in the study of Zenner et al. (2012) with the same length as the English loanword, we conducted a pretest with a small group of linguistic students who were asked to indicate the most suitable near-synonym for the Anglicism. This leaves a set of Anglicisms and Dutch alternatives where speech economy cannot factor into respondents' selection (cf. Winter-Froemel et al. [2014] and Schaefer [2019] for the influence of word length on the usage of Anglicisms). Second, to arrive at a representative selection of Anglicisms in Dutch, we tried to balance the stimuli in terms of the semantic field, concept frequency, and corpus-based success rate of the Anglicisms (see Zenner et al. 2012 for details). The concepts belonged to three semantic fields which often include Anglicisms: sports, jobs, and media (see above). Concerning concept frequency, i.e., the number of times a concept occurred in the corpus, two concepts were highly frequent, seven moderately and three appeared only a limited number of times in the newspaper corpora. Finally, concerning the corpus success rate, we deliberately chose to include both

4 Thus, if we take the abovementioned instance of *keeper* (containing two syllables), only the Dutch alternative *doelman* (containing two syllables as well) could be selected, as the other heritage forms have either a length of less or more than two syllables.

concepts for which the heritage forms had a very low amount of occurrence (and whose borrowed counterparts hence had a high success rate) and concepts for which the heritage form occurred frequently. This allowed checking the influence of community-based usage patterns. As a result, the definite set of stimuli (see Table 1) approaches the idea of a golden mean mentioned above, as it allows a careful investigation of several variables over multiple Anglicisms.

Besides the 12 English – Dutch word pairs, 10 filler items were included in the experiment. For the most part, these consisted of Dutch word pairs. However, some highly established loanwords from French such as *auteur*, *artiest*, *crimineel,* and *arbiter* were included, as well as the internationalism *professor* (see Table 2). This was done to distract the participants from recognizing that the goal of the experiment was to investigate English – Dutch oppositions.

4.2.2 Sanity-checks: (near-)synonymy and English loanword status

The second part of the questionnaire consisted of two sanity checks. The first sanity check was introduced to test whether the participants reckoned the pairs to be near-synonyms, as we wanted each word pair to consist of near-equal opponents (see Edmonds and Hirst 2002). We specifically asked "To what extent do the following words mean the same to you?" to which the participants could answer using a five-point Likert scale (1 = not at all the same, 5 = completely the same). Besides the 12 trial targets from the forced-choice experiment, six fillers (four from the forced-choice test and two new word pairs) were added to help benchmark the scales. Some fillers were, in our opinion, clearly synonymous and would receive very high scores (e.g., *dokter* and *arts*, "doctor" and "physician"), while others were not and would help benchmark the right side of the scale (e.g., *basketballer* and *voetballer*, "basketball player" and "soccer player"). A second sanity check was included to assess whether the informants recognized the Anglicisms to be English. They had to indicate on a five-point Likert scale to what extent the words felt like English loanwords (1 = feels like a purely English word, 5 = feels like a purely Dutch word). Again, fillers were added to benchmark the scales, e.g., *account executive,* which is certainly an English word, and *leerkracht,* which is certainly a Dutch word.

4.2.3 Respondents' relationship with English

Part three of the survey openly questioned the level of contact with English. First, respondents were asked about their own use of English in terms of writing and speaking (daily, weekly, monthly, less than once a month). Next, exposure to English was gauged by asking the participants how frequently they read and heard

Table 1: Set of stimuli.

Anglicism	Heritage form	Semantic field	Concept frequency	Corpus-based success rate Anglicism	Corpus-based success rate Dutch term	Time of borrowing
keeper	*doelman*	sports	189,533	17.20%	82.80%	before 1989
pitcher	*werper*	sports	1,515	56.50%	43.50%	before 1989
jockey	*rijknecht*	sports	972	99.79%	0.21%	before 1989
catcher	*vanger*	sports	184	98.91%	1.09%	before 1989
covergirl	*hoezenpoes*	media	129	87.60%	12.40%	before 1989
songwriter	*liedschrijver*	media	1,832	42.74%	57.26%	before 1989
ghostwriter	*spookschrijver*	media	380	99.47%	0.53%	before 1989
quizmaster	*quizleider*	media	1,114	97.85%	2.15%	before 1989
businessman	*zakenman*	jobs	28,675	0.47%	99.53%	before 1989
marketeer	*marktstrateeg*	jobs	2,023	94.56%	5.44%	before 1989
babysitter	*kinderoppas*	jobs	1,18	36.78%	63.22%	before 1989
workaholic	*werkverslaafde*	jobs	1,207	97.76%	2.24%	before 1989

Table 2: Set of fillers.

Fillers	
baas	*gezagvoerder*
professor	*onderwijzer*
kunstenaar	*artiest*
eigenaar	*bezitter*
crimineel	*misdadiger*
schoonmaakster	*poetsvrouw*
landbouwer	*boer*
scheidsrechter	*arbiter*
pedagoog	*opvoeder*
auteur	*schrijver*

English (daily, weekly, monthly, less than once a month). Additionally, they had to describe their level of English proficiency (below average, average, good, expert). Note that not enough variation was attested between respondents to incorporate these parameters as a predictor in the analyses, though see Section 4.4 for information on respondents excluded from the final sample, based on proficiency issues. Finally, to answer RQ 2.4, we asked respondents whether the use of English bothered them considering their own self-reported usage:

1. *Ik stoor me er niet aan, ik gebruik zelf ook Engelse woordenschat.*
   "It doesn't bother me, I use English vocabulary myself."

2. *Ik stoor me er niet aan, al gebruik ik zelf weinig tot geen Engelse woordenschat.*
   "It doesn't bother me, although I hardly use any English vocabulary myself."

3. *Ik heb hier geen mening over. Ik let er niet op.*
   "I don't have an opinion about this. I do not pay attention to the use of English words."

4. *Ik stoor me er aan, al gebruik ik zelf ook Engelse woordenschat.*
   "It does bother me, although I use English vocabulary myself."

5. *Ik stoor me er aan, ik gebruik dan zelf ook weinig tot geen Engelse woordenschat.*
   "It does bother me, so I hardly use any English vocabulary myself."

Finally, participants were invited to freely write down any further comments they had on the use of English.

4.2.4 Socio-demographic information

The final part of the survey included questions about the socio-demographic characteristics of the respondents. Specifically, we asked about age, gender, nationality, mother tongue, and their level of formal education.

4.3 Procedure

As mentioned above, the forced-choice experiment was embedded in an online, anonymous four-part survey, which was distributed using the social media platforms Twitter and Facebook calling "people with an opinion." Besides sharing the survey on our personal profiles, we also posted it on Dutch and Flemish Facebook groups with a large number of members to (i) attract informants from both nationalities and (ii) to bring in as many participants as possible. While we did not say that the goal was to investigate language attitudes toward English loanwords, we did inform the participants that we needed their opinion on certain words. Correspondingly, the title of the survey was: "Choosing between words: What does your *taalgevoel* 'language intuition' tell you?"

The first page of the questionnaire informed the participants that this study was conducted in the context of an ongoing research project investigating the lexical choices of Dutch native speakers from Flanders and the Netherlands and that the survey was completely anonymous. They were further told they would be presented with an experiment requiring them to choose between synonyms 22 times. Starting on the second page, the forced-choice experiment presented each question on to a page. Following the forced-choice experiment, participants were first to fill out the synonymy sanity check, then the loanword status check. This time all word pairs were shown together. The next page contained five questions concerning the participants' relationship with English (production, exposure, English level of proficiency, self-reported attitude, and the comment section). Finally, the last page contained the socio-demographic questions about age, gender, nationality, mother tongue, and level of formal education. Once the participants left a page, they could not return to it again. This was done to prevent participants from going back to the forced-choice task and changing their answers, once they realized the study was about the use of English words. In total, the survey took about 15 min to complete.

4.4 Respondents

In total, we received 1,500 completed responses, yet only 1,228 surveys were retained for further analyses. The following participants (N = 272) were excluded:

- participants who did not provide information on their year of birth or gender;
- participants who are under 18 or over 70 years old, as the sample was too small in these age ranges to allow for thorough analyses;
- participants who did not identify with male or female gender, again as the number of respondents in this category was too low to allow for further scrutiny;
- for the same reason, respondents without a secondary school degree or who claimed to have no reasonable command of or exposure to the English language;
- participants who are not Dutch monolinguals;
- participants who provided clearly unreliable answers in the sanity checks or who used the option "I don't know (one of these) words" in more than 4 of the 12 target trials. Note that several of the excluded surveys were filled out by participants who met more than one of these criteria for exclusion.

The relevant socio-demographic information on the 1,228 retained surveys is found in Table 3. We categorized respondents into three age groups: group 1: 18 to 29 (student population and early starters, whom Bontekoning 2007 labels *juniors*), group 2: 30 to 50 (active population with active family life, whom Bontekoning defines as *mediors*), and group 3: 51 to 70 (Bontekonings' *leaders*). For self-reported attitude toward English, we relied on a binary reinterpretation of the variable to avoid issues with data sparseness. Overall, the table reveals several skews in the sample, yet also shows that we still have a reasonable signal

Table 3: Socio-demographic information on the respondent sample.

		N	%
Age (RQ 2.1)	Group 1: 18–29	666	54%
	Group 2: 30–50	349	28%
	Group 3: 51–70	213	18%
Gender (RQ 2.2)	F	945	77%
	M	283	23%
Region (RQ 2.3)	Flanders (Belgian Dutch)	832	68%
	The Netherlands (Netherlandic Dutch)	396	32%
Self-reported attitude (RQ 2.4)	The use of English vocabulary bothers me	281	23%
	The use of English vocabulary does not bother me + no opinion	947	77%

in the different cells. However, combined with the repeated measures for the respondents and trials, this type of skew means that advanced inferential statistical techniques are required to analyze the results of our experiment.

4.5 Analysis

The research questions are tackled by employing mixed-effects logistic regression analyses. The response variable is binary, contrasting the selection of the English word with a selection of the Dutch word. The wildcard option from the survey ("I don't know [some of these] words") is considered an indication that the respondent cannot make a reliable choice between the English word and the Dutch alternative and is hence discarded from the analyses (though see more information on this answer option in Section 5.2). In light of the research questions, six predictors are included in the model selection procedure: Corpus success of the Anglicisms (based on Zenner et al. 2012), Age of the respondent (group 1: 18–29; group 2: 30–50; group 3: 51–70), Gender of the respondent (female; male), Region (Flanders; the Netherlands), Reported attitude toward English influence (bothered; not bothered or neutral), and Context offered in the trial (newspaper article; conversation with friends). Both the respondent ID and the word pair under scrutiny are included in the model as a random effect, as both come with repeated measures: each respondent runs through 12 trials, and each word pair is assessed by 1,228 respondents. Factoring in the potential auto-correlation for respondents and words when estimating the contribution of the fixed effects to explaining the attested variation provides a first justification for the selection of mixed-effect regression analyses. Additionally, regression analyses allow us to assess the behavior of the predictors while keeping the effect of other predictors stable, which is particularly crucial with the type of skewed samples (see Table 3) and more complicated mixed designs (e.g., the random distribution of context per trial per respondent) we work with here. More information on the model building process, the goodness of fit, and model interpretation will be presented in Section 5.2.

5 Results

Before addressing the results of the mixed-effect regression model, we report the results of the two sanity checks in Section 5.1, complemented with a check of the overall variation for the target items. Based on these sanity checks, two of the 12 target stimuli will be discarded from further analyses that will be discussed in Section 5.2.

5.1 Verifying the guiding assumptions in stimuli selection

Two core assumptions guided our selection of target stimuli (the alternation pairs of English lexemes and Dutch alternatives), viz. (1) that the two words were considered semantic equivalents; (2) that the English words were perceived as English loanwords, or at least as not equally "Dutch" as the heritage terms. The second part of our survey verified to what extent these assumptions hold for our respondent sample (see Section 4.2.2 for more details).

Figure 1 reports the results of the first sanity check, where respondents were asked to rate the semantic equivalence of word pairs on a five-point Likert scale (1 = no similar meaning/not felt to be synonymous, 5 = similar meaning), including not only the 12 target word pairs but also six fillers pairs that were meant to benchmark the rating scale (*architect_bouwvakker* "architect–construction worker," *basketballer_voetballer* "basketball player–football player," *professor_onderwijzer* "professor–teacher," *landbouwer_boer* "farmer–peasant," *arts_dokter* "physician–doctor," *auteur_schrijver* "author–writer," see Section 4.2.2). First, the position of the filler items in Figure 1 (white boxplots) reveals a well-benchmarked scale. Second, the overall results for the target items (grey boxplots) are reassuring. Only three word pairs have a mean synonymy rating under 4 on the 5-point scale,

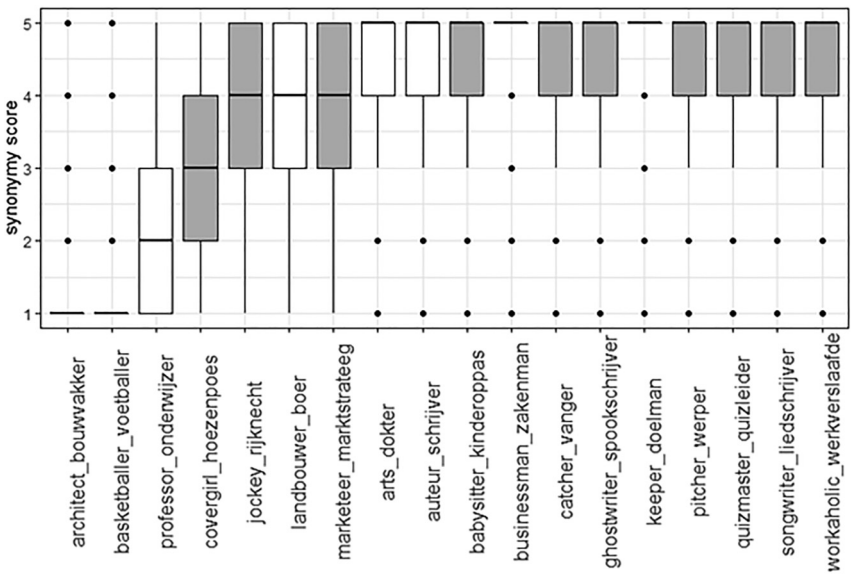

Figure 1: Sanity check 1: loanword status score (1 = not perceived as similar in meaning, 5 = perceived as similar in meaning); target items in grey, filler items in white.

viz. the word pairs *covergirl – hoezenpoes* (M 3.20, SD 1.34), *jockey – rijknecht* (M 3.73, SD 1.25), and *marketeer – marktstrateeg* (M 3.77, SD 1.11). Particularly the boxplot for *covergirl – hoezenpoes* reveals a problematic sanity check, forcing the removal of the item from further scrutiny.

The results for the second sanity check, which gauges the perceived loanword status of the English items (1 = felt to be an English loanword, 5 = felt to be part of the Dutch lexicon), are reported in Figure 2. The filler items that are meant to help benchmark the scale are found in white. The boxplots for the target items are found in grey. Again, the filler items reveal an intuitive scale, with foreign words such as *police officer* or *secretary of state* receiving very low scores, as opposed to entrenched loanwords included in Dutch dictionaries such as *ballerina* or *architect,* and Dutch items such as *leerkracht* "teacher" or *verkoopster* "seller," which receive high scores. The results for the 12 English target loanwords are in this case also reassuring: none of the 12 items has a mean foreignness score under 2, and each of the 12 items has a median foreignness score of 1.

As a final check, we had a look at the answer patterns for each of the 12 stimuli. Before addressing the core research question on the variation in lexical selection of Anglicisms and heritage alternatives, we wanted to make sure that there is, as we assumed, effectively notable variation in the answer patterns for the loanwords.

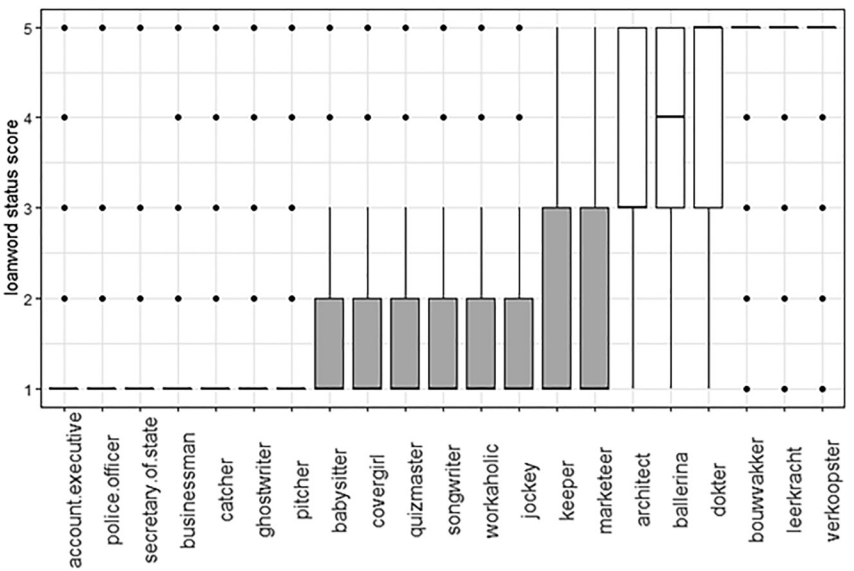

Figure 2: Sanity check 2: loanword status score (1 = perceived as foreign, 5 = perceived as native); target items in grey, filler items in white.

As discussed in Section 4.2.1, we deliberately included some alternation pairs of which the corpus-based success of the Anglicism was very high (e.g., *covergirl – hoezenpoes*). We included these items to gauge whether respondents' purist reflex might override the usage-based nature of lexical choice, yet we herewith introduce the risk that a sizeable proportion of respondents might not know the Dutch alternative and decide to refrain from lexical choice by selecting the wildcard "I don't know (one of) these words". Table 4 reveals the number of times respondents selected this wildcard option for each of the words, complemented with information on the number of respondents who selected the English or Dutch word per trial. Finally and crucially, the table also provides the proportion of respondents who selected the English item when excluding respondents who withdrew from selection through the wildcard, revealing the amount of variation attested in the answer patterns. The table reveals that two items are not fit for further analyses: both *jockey_rijknecht* en *covergirl_hoezenpoes* are the only items (1) where more than 10% of respondents selected the wildcard option "I don't know (one of) these words"; (2) where barely any variation in answer patterns is left when excluding these wildcard answers, with more than 95% of respondents selecting the English term for both alternation pairs. For *covergirl – hoezenpoes*, this adds to the already problematic results from the first sanity check.

To conclude our verification of the guiding assumptions underlying the stimuli selection, two-word pairs are excluded from further analyses: *jockey – rijknecht* and *covergirl – hoezenpoes*. This leaves us with 11,892 forced choices for 10 target alternation pairs obtained from 1,228 respondents. In total, the English term

Table 4: Variation in answer structure per trial.

Target trial	*N* wildcard	*N* English	*N* Dutch	(*N* English)/ (*N* English + *N* Dutch)
jockey – rijknecht	147	1,075	6	0.994
covergirl – hoezenpoes	123	1,088	17	0.985
quizmaster – quizleider	1	1,164	63	0.949
ghostwriter – spookschrijver	57	1,109	62	0.947
songwriter – liedschrijver	2	1,149	77	0.937
workaholic – werkverslaafde	0	1,104	124	0.899
marketeer – marktstrateeg	86	910	232	0.797
pitcher – werper	121	853	254	0.771
keeper – doelman	0	921	307	0.750
catcher – vanger	121	761	346	0.687
babysitter – kinderopvang	0	829	399	0.675
businessman – zakenman	0	197	1,031	0.160

was selected in 8,997 trials (75.7%), the Dutch term was selected in 2,895 trials (24.3%). In the next section, we attempt to increase our understanding of what guided respondents to make their selection.

5.2 Addressing the research questions

As explained in Section 4.5 above, we analyze the data using mixed-effects logistic regression analyses. As concerns the response variable, we estimate the chance of selecting the Dutch item (1) over the English item (0). Although this might seem conceptually counterintuitive, it is technically the most robust way to fit the model, considering the reasonable skew in the response (75.7% English, 24.3% Dutch). In terms of the model-building strategy, we started with a null model including only random intercepts for the variables "Respondent ID" and "Stimulus (word pair)". The fixed effects were included through a manual forward stepwise selection relying on AIC, confirmed through Chi-square tests that compare the null model and the fitted model. The best-fitting model includes four of our six variables, viz. "Corpus success of the Anglicism" (see Table 1),[5] "Context" (newspaper article or conversation with friends), "Age of the respondent" (group 1: 18–29; group 2: 30–50; group 3: 51–70), and "Self-reported attitude respondent" (bothered vs. not bothered/neutral).

In terms of fit, the model reported in Table 5 (random effects) and Table 6 (fixed effects) significantly reduces the null model's AIC (from 9,958.580 to 9,749.5, for 11,884 residual degrees of freedom in the fitted model). The relative contribution of the different fixed effects to explaining the variation in the response variable is presented in Table 7, with higher Chi-square values indicative of a stronger contribution to explaining the variation in the response. In terms of power, we find a C-measure of 0.878. This measure has values between 0 and 1, with values over 0.8 seen as indicative of models with predictive power. The conditional R^2 of the model is 0.463. This value between 0 and 1 indicates the amount of variation in the response variable explained by the model. Where both the C-measure and the conditional R^2 are calculated taking both fixed and random effects into account, the marginal R^2 of the model, a value between 0 and 1, indicates how much of the variation in the response variable is captured exclusively by the fixed effects in the model (see Winter 2019: 264). This model achieves a reasonable marginal R^2 of 0.22.

5 The model includes a transformation of the variable: instead of working with the relative success rates of the Anglicism vis-à-vis the Dutch alternative, we include the log(odds) of the success for a better model fit.

Table 5: Information on the random effects in the mixed-effects model.

		Variance	Standard deviation
Respondent ID (N = 1,228)	(Intercept)	0.530	0.728
Stimulus (Word pair) (N = 10)	(Intercept)	0.965	0.9821

Table 6: Information on the fixed effects in the mixed-effects model (estimating the chance for a respondent preferring the Dutch lexicalization).

	Estimate	Std.Error	*Z*-value	Pr(>|*z*|)	
(Intercept)	−0.944	0.339	−2.782	<0.01	**
Corpus success of the Anglicism	−0.339	0.010	−3.447	<0.001	***
Context ("Friends" vs. "Newspaper")	−0.664	0.054	−12.255	<0.0001	***
Age respondent ("group 2: 30 to 50" vs. "group 1: 18 to 29")	−0.195	0.079	−2.478	0.01	*
("group 3: 51 to 70" vs. "group 1: 18 to 29")	0.263	0.091	2.901	<0.01	**
Self-reported attitude respondent (bothered vs. not bothered/neutral)	0.505	0.078	6.437	<0.0001	***

Table 7: Analysis of deviance table (Type II Wald chi-square tests), revealing the relative contribution of the predictors to explaining variation in the model.

	Chisq	*df*	Pr(>Chisq)	
Corpus success Anglicism	11.88	1	<0.01	**
Context	150.18	1	<0.0001	***
Age respondent	20.75	2	<0.0001	***
Self-reported attitude respondent	41.44	1	<0.0001	***

For our interpretation of the model, we focus on the fixed effects (Table 6); the random effect structure will not be discussed further. Overall, the most relevant information for addressing the research questions is captured in the second and the final column of Table 6. The final column of the table provides information on the significance of the effects, with $p < 0.05$ indicating significance, and more stars indicating lower p-values. For significant effects, the next step is to scrutinize the second column. This column includes the model's estimates for the effects, with positive estimates indicating a higher chance of respondents selecting the Dutch

item and negative estimates indicating a higher chance of respondents selecting the English item. Below, we make this more concrete for each of the parameters, addressing our research questions one by one.

To answer the first research question, we need to verify whether the community-based usage statistics for the word pairs under scrutiny help explain the respondents' self-reported preference for English loanwords or their Dutch alternatives in the experiment. Table 6 reveals a significant effect for the predictor "Corpus success Anglicism," although Table 7 indicates that the parameter has the weakest contribution of the four selected predictors to the model: we see a significant effect with a smaller effect size. For the interpretation of the effect, we turn to Figure 3, which presents us with the numeric predictor "Corpus success Anglicism" (expressed as log(odds)) on the *x*-axis and the model's fitted probability of respondents selecting the Dutch term in the trials on the *y*-axis: the higher the value on the *x*-axis, the more likely respondents are to select the Dutch term over the English term. We can clearly see the anticipated negative trend: as the corpus success of the Anglicism goes up, the probability of the respondent selecting the Dutch term goes down. We can hence favorably answer RQ1: the community-based usage statistics for the word pairs under scrutiny help explain the respondents' self-reported preference for English loanwords or their Dutch

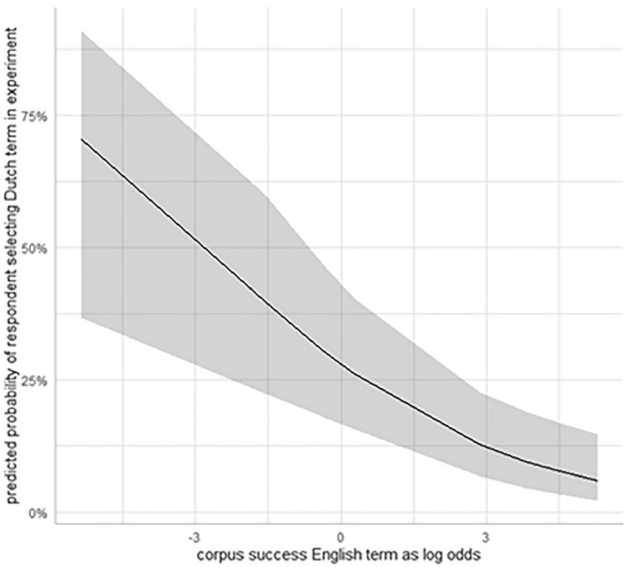

Figure 3: Effect plot corpus success English term.

alternatives in the experiment, with reported preference resonating usage patterns.

For the second research question, viz. "Do the socio-demographic characteristics of the respondent help explain the respondents' self-reported preference for English loanwords or their Dutch alternatives in the experiment?", the effect of four parameters was tested in the model, viz. "Age," "Gender," and "Region" of the respondents and their "Self-reported attitude" toward English influence. The two parameters for which less straightforward hypotheses were presented – "Region" and "Gender" – do not reach significance in the model. The age of the respondent and their self-reported attitude do play a role. For "Age," the regression model captures the behavior of the youngest respondents (group 1: 18–29) in the intercept, then compares the behavior of the two other age groups to this intercept. Both comparisons yield significant differences: we see a significantly lower estimate for the odds of selecting the Dutch term for "group 2: 30–50," than for the youngest participants (group 1: 18–29) and a significantly higher estimate for the odds of selecting the Dutch term for "group 3: 51–70," than for the youngest participants (group 1: 18–29). Put differently: it is not as anticipated the group of the youngest respondents, but it is the active population (group 2: 30–50) who is most likely to select the English lexeme. Congruent with our hypothesis, the oldest age group (group 3: 51–70) is least likely to select the English lexeme. The results for "Self-reported attitude" toward English are highly intuitive: respondents who are allegedly bothered by the use of English have a higher chance of selecting the Dutch term in our experiment than respondents who claim not to be bothered by English influence or who state not to have an opinion (captured in the intercept). We can thus favorably answer RQ2: although not all of the variables included have an effect and although not all effects confirm the specific hypotheses introduced, we do see that some of the socio-demographic characteristics of the respondents help explain their lexical choices for English loanwords or their Dutch alternatives in the experiment.

Finally, we turn to RQ3: "Do speech contexts of communicative immediacy versus distance (conversation with friends or quality newspaper article) help explain the respondents' self-reported preference for English loanwords or their Dutch alternatives in the experiment?". Firstly, this parameter has a notably larger impact on explaining the variation than the other three significant predictors (see Table 7). Next, studying the effect, we see that the probability of respondents selecting the Dutch term in trials is lower in the context of "conversation with friends" than in the context of "newspaper article." English terms are allegedly considered more appropriate in the immediacy context of casual conversation than in the distance context of the newspaper article. Hence, we can favorably answer RQ3: the speech context introduced to the respondents in the experimental trial

helps explain the respondents' lexical preference for English loanwords or their Dutch alternatives in the experiment, and the parameter of communicative immediacy versus distance plays an important role for the lexical choices made. In the next section, we will unpack these results with some more attention to their theoretical implications, meanwhile attenuating the analyses with some methodological reflections where needed.

6 Discussion

From a theoretical perspective, it is interesting to dig deeper into four results introduced in Section 5: (1) the correlation between community-based usage statistics and the reported lexical preferences for the studied word pairs, (2) the non-linear effect of age on the studied lexical preferences, (3) the lack of a region effect on the reported preferences and (4) the overriding effect of speech context as a predictor in respondents' lexical choices.

While it might not come as a surprise that loanwords that occur more frequently in a newspaper corpus are also more likely to be selected as "most appropriate" in a lexical-preference experiment, the correlation is interesting to highlight, as it aligns nicely with a usage-based perspective on language variation. Usage-based approaches to language assume that (1) linguistic knowledge is an inventory of cognitive "units" – form-meaning combinations of varying schematicity – in an individual's memory and that (2) this inventory results from and is constantly updated based on the usage events an individual engages in. Hence, a usage-based approach expects a clear link between language usage on the one hand and the mental representation an individual has of language usage on the other. In the case of Anglicisms, the frequency of occurrence can be expected to correlate with the loanword's "entrenchment" in cognition, which in turn determines how appropriate the Anglicism is deemed, and also influences new usage events. Indeed, our results show a significant, proportional effect of an Anglicism's frequency in a newspaper corpus on lexical preferences. However, the correlation between corpus success and the probability that a respondent selected an English loanword is not absolute. The regression model shows a significant, but rather small effect of corpus frequency. In our view, two factors seem to be at play here.

Firstly, attitudinal factors can explain why the reported preferences do not perfectly reflect corpus success frequencies. Linguistic purism, for instance, i.e., the desire "to preserve a language from, or rid it of, putative foreign elements, or other elements held to be undesirable" (Thomas 1991: 12), might influence

reported preferences more than actual language usage. This was explicitly suggested by respondent 413 in the open comments section of our survey:

(1) *Hoewel er vaak gedacht wordt dat de integratie van Engelse woordenschat in het Nederlands leidt tot een zogezegde "verloedering" van de taal, vind ik dat er een paradox is tussen het feit dat zoooo veel mensen hierin geloven maar toch dagelijks Engelse woordjes in hun taal gebruiken.*
"Though it is often thought that integration of English vocabulary in Dutch leads to a so-called 'corruption' of the language, I think that there is a paradox between the fact that soooo many people believe this, but still use English words daily in their language."

Puristic reasoning also clearly emerges in other comments, mainly "metalinguistically" and "ideologically" motivated (cf. Langer and Nesse 2012). Langer and Nesse (2012: 610–611) identify – following Gardt (2001) – four types of discourse underlying puristic reasoning: (1) *structural discourse,* which "assumes that for each language there is a state of purity at which the linguistic system is perfectly balanced", (2) *ideological discourse,* which "emphasizes the superior quality of one's language based on its genealogical purity and great age" and which views foreign elements "as a corruption of cultural purity", (3) *pedagogical discourse,* which argues "that the use of foreign words leads to social division within a society, since the less-educated and the elderly might not understand new borrowings introduced into specific domains such as youth language or technical registers" and (4) *metalinguistic discourse,* which "more openly acknowledges that purism has to do with taste and aesthetics, rather than a general ability to communicate" and scorns the use of foreign words as "chasing fashionable trends", "giving the impression of being intellectual and modern" and hence also as "superficial and pretentious behavior". In our survey, 87 respondents (7.1%) made use of the open comments section, of which 4 respondents criticized English loanwords based on ideological arguments (cf. Examples (2) and (3)) and 5 based on metalinguistic ones (cf. Example 4). Structural or pedagogical arguments were not given explicitly, though reading between the lines, e.g., in Examples (3) or (4), structural or pedagogical reasoning can also be found. The other comments mainly pertained to the personal linguistic background of the respondents (n = 12), the research design (n = 9), indicated an open attitude toward English (n = 19), or commented on the difference between "necessary" and "luxury" loans (n = 15). We refrained from verifying statistically whether there is a correlation between the answers in the comments section and the answers in the forced-choice experiment, i.e., whether participants who commented negatively on the use of English in the open answer field choose fewer Anglicisms and vice versa. We did so because, on the one hand, not all the answers allow for a crude categorization, and on the

other, because not enough participants made use of the open answer field. Nevertheless, it would be interesting to investigate this in further research.

(2) Resp. 178: *Het Nederlands ís een rijke taal, het is absoluut geen schande er de voorkeur aan te geven.*
"Dutch is a rich language, it is absolutely no disgrace to favor it."

(3) Resp. 255: *Jammer dat wij meer dan eeuw letterlijk hebben gevochten om het Nedrlands* (sic) *als eenvolwaardige* (sic) *taal te laten herkennen, in eerste instantie tov het Frans om nu onze taal volledig in de verdomhoek te gooien voor hetEngles* (sic).
"It's a shame that we have literally fought for more than a century to have Dutch recognized as a respectable language, initially vis-à-vis French to now just throw the language completely away for English."

(4) Resp. 453: *Het stoort me vooral als het Nederlands een uitstekend synoniem heeft, en al helemaal als dat Nederlandse synoniem ook nog eens korter én bekender is. Of als ik de indruk krijg dat de spreker/schrijver de woorden gebruikt om interessant te doen, of omdat hij/zij niet eens weet dat ze hetzelfde betekenen als een bestaand Nederlands woord. Maar omgekeerd moeten we ook niet alles willen vertalen – ik las laatst een tekst na voor een collega en daarin werd "empowered" vertaald als het neo-hippie-klinkende "in haar kracht staand".*
"It mainly bothers me when Dutch has an excellent synonym, and certainly when that Dutch synonym is also shorter and more well-known. Or when I get the impression that the speaker/writer uses the words to be interesting, or because he/she does not even know that they mean the same as an existing Dutch word. But conversely, we also do not have to want to translate everything – I recently revised a text for a colleague in which "empowered" was translated as the neo-hippy-sounding 'in haar kracht stand'."

Attitudes and beliefs influence the cognitive representations of linguistic items and can explain why reported preferences do not perfectly reflect corpus success frequencies. It has been shown that there is a general tendency for highly salient innovations to frequently give rise to strong judgmental reactions by the speakers (Winter-Froemel 2020), and it could thus be expected that these factors will have an impact on the reported lexical preferences for highly frequent and highly salient Anglicisms. On the other hand, it has often been observed (see e.g., Kabatek 1996) that the speakers' actual linguistic behavior does not necessarily match and may even contradict their self-reported linguistic attitudes. A complex interaction of

factors thus needs to be assumed, and this should of course be accounted for in an adequate usage-based model on language variation (see e.g., Campbell-Kibler's 2016 proposal of a "cognitively realistic model of sociolinguistic variation," in which she assumes the existence of a self-regulation system, which allows language users to control their speech production, perception, and attitudes, but would be limited in the number and complexity of constructs it can monitor).

Another factor potentially explaining the weak effect size of corpus success is that the frequency data from Zenner et al. (2012) do not perfectly mirror the frequencies with which the respondents encounter the studied Anglicisms in real life. The entrenchment of linguistic units varies among and within individuals (cf. Verhagen 2019), and hence, the entrenchment of the loanwords for the individual respondents does not completely match the "community entrenchment" measured using frequencies in a newspaper corpus.

Variation in entrenchment might not only account for the weak correlation between the corpus success measured in Zenner et al. (2012) and the reported lexical preferences, but also for the age effects in our data (see also Walsh 2015). Recall from Section 5 that the active population (group 2: 30–50) in our study was most likely to select English lexemes – more than the youngest generation (group 1: 18–29) – and that the oldest age group (group 3: 51–70) was least likely to do so. Our assumption here is that these differences correlate with differences in the degree to which members of the studied age groups encounter the presented Anglicisms. Research by Zenner et al. (2013) has shown an increase in English loanwords in Dutch job advertisements between 1995 and 2008, and hence, interpreted in line with the apparent-time hypothesis, which argues that ongoing language change is reflected in differences among age groups in the present (cf. Labov 1963), the entrenchment of English loanwords can be assumed to be weaker among older language users. Given that all word pairs presented to our test subjects included Anglicisms borrowed in Dutch before 1989, this reasoning seems not to hold quite well though. The relatively old Anglicisms included here could rather help explain why, contrary to our expectations, the youngest generation (group 1: 18–29) was not more prone to select the Anglicisms than the "active population."

Similar familiarity-based reasoning can be applied to explain the lack of region effects in our data. Our study hypothesized that Belgian Dutch participants would react with stronger purist anti-English reflexes to the stimuli than Netherlandic Dutch respondents, given the standardization history of Dutch in Flanders (cf. Section 1). This idea was also voiced by some of our respondents. One respondent for instance remarked the following in the open comments section:

(5) Resp. 144: *In Nederland wordt* (sic) *er veel meer Engelse woorden gebruikt in de dagelijkse omgang dan in Vlaanderen tijdens een gesprek over dezelfde onderwerpen.*
"In the Netherlands, more English words are used in daily interaction than in Flanders in conversations about the same subjects."

However, in the lexical preferences themselves, there was no difference between Netherlandic Dutch respondents and Flemish respondents. This might again have to do with the selected Anglicisms. The Anglicisms were selected from the corpus study of Zenner et al. (2012), which also did not report significant differences between the Flemish and Netherlandic subcorpora. At the community level, the presented Anglicisms seem to be entrenched equally in both the Netherlands and Flanders, and it is hence not surprising that the individual appropriateness judgments do not differ significantly. Yet, as we highlighted at the beginning of this section, the correlation between newspaper data and the individual choices reported in this study is far from absolute, so there might be other factors involved. It is also possible that the region effect we hypothesized does not exist in production, but only in attitudes. In this scenario, the reported attitudes of the Belgian Dutch and Netherlandic Dutch participants would be different. This proves to be the case, though contrary to expectations we find a higher proportion of Netherlandic Dutch participants who claim to be bothered by the use of English (117 out of 396, viz. 29.5%) than in the Belgian Dutch sample (164 out of 832, viz. 19.7%). The association between region and self-reported attitude is significant ($X2$ $(1, N = 1{,}228) = 14.15$, $p < 0.0001$), but fairly weak (Cramer's V 0.1); more research is needed to grasp this pattern.[6]

A last and interpretatively more straightforward result to highlight is the importance of context in our experiment. Context was the most important predictor of the success of an Anglicism in our study: the probability of respondents selecting an English term was significantly higher in the context of "conversation with friends" than in the context of "quality newspaper article." English loanwords are thus considered more appropriate in the immediacy context than in the

6 An anonymous reviewer interestingly suggested that the higher level of societal multilingualism in Belgium might play a role in these findings, as participants with multilingual repertoires might have a higher acceptance of foreign items. However, the three national languages in Belgium each correspond with an official territory (except in Brussels, which is officially bilingual). Thus, we expect the level of individual multilingualism of Belgian speakers to be quite similar to Dutch speakers, as is also supported by the data. Additionally, as indicated in Section 4.4, we only retained monolingual Dutch speakers, viz. speakers who reported to only have Dutch as native language. Nevertheless, further research is needed to grasp the link between societal/individual multilingualism and the attitudes to foreign languages.

distance context. The importance of context was also stressed by 15 respondents in the open comments (cf. Examples (6) and (7), with the respondent in Example (7) explicitly pointing to the importance of the written realization, which is typical for communicative distance).

(6) Resp. 110: *Ik zou graag de nuance toevoegen dat mijn irritatie aan de Engelse influx vooral afhankelijk is van de bron: wanneer een kwaliteitskrant of een andere hoogstaande, verder strikt nederlandstalige bron zich spontaan niet meer voldoende in de moedertaal uit kan drukken en kennelijk uit lijkt te moeten wijken naar het Engels, zie ik dat als een tekortkoming. Wanneer een willekeurig social mediabericht of een flutartikeltje van een verder niet relevante bron zich eraan vergrijpt, vind ik het geen enkel probleem.*
"I would like to add the nuance that my irritation at the English influx mainly depends on the source: if a quality newspaper or another high-minded, otherwise strictly Dutch source can no longer express itself spontaneously in the mother tongue, and seemingly cannot but recourse to English, that is a shortcoming to me. If a random social media comment or a crummy article of an otherwise not relevant source does the same, I don't mind at all."

(7) Resp. 709: *Ik vind het gebruik van Engelse leenwoorden in de omgang geen probleem, maar in de geschreven taal kan ik me er wel degelijk aan storen.*
"I don't mind the use of English loanwords in interaction, but it can bother me in written language."

The results suggest that contexts of communicative immediacy or distance are of key importance for the selection of Anglicisms or their Dutch alternatives. This observation potentially has far-reaching theoretical implications, as it means that speakers do not have a uniform evaluation of the Anglicisms, but strongly take into account the specific context of the use of the items, which may lead to different evaluations of the adequacy of the same Anglicism by the same speaker. From a methodological perspective, the results stress the importance of incorporating context into attitudinal experiments (cf. Nejjari 2020; Rosseel et al. 2019). At the same time, it might be the case that the effect of context interacts with the level of entrenchment of the Anglicisms. In this study, we only included well-entrenched loanwords that were borrowed before 1989. However, it is possible that more recent Anglicisms, as well as Anglicisms belonging to other registers and semantic domains, might be evaluated in another way. Therefore, follow-up studies should include a variety of Anglicisms borrowed from different periods covering other semantic domains and registers to obtain a more fine-grained picture of the pragmatic factors steering the lexical choice of speakers.

7 Conclusion

The present study aimed at contributing to Anglicism research by studying the evaluation of Anglicisms experimentally and from an onomasiological perspective. By focusing on lexical preferences of Anglicisms compared to heritage alternatives, the study was able to show that lexical preference is steered by a combination of sociopragmatic parameters, with the context of usage playing a crucial role in the speakers' self-reported usage. Of course, there are a few methodological concerns to bear in mind. First, replication of this study with more alternation pairs would be a valuable addition to the present study, preferably also integrating loanwords that are less entrenched (e.g., incipient loanwords or single-word switches). Secondly, it might be interesting to complement the current design with an experiment that is more indirect about its purpose. The here presented study did integrate fillers to divert the attention somewhat from the Anglicisms, but of course, the goal of the research was still very clear. To conclude, the "sanity checks" of our design indicated heterogeneity in the degree to which the presented Anglicisms were perceived as "foreign." This issue was only touched upon briefly in our report – we mainly attempted to keep the degree of synonymy and perceived exogeneity constant – but it would be interesting to investigate why some words are more easily experienced as "foreign" than others. There are in other words still a lot of intriguing questions to tackle, but we do hope to have demonstrated the usefulness of an experimental, attitude-oriented approach toward lexical choice and to have paved the way for further research expanding on the theoretical and methodological implications of the findings obtained.

References

Ager, Dennis Ernest. 1999. *Identity, insecurity and image: France and language*. Clevedon: Multilingual Matters.

Andersen, Gisle. 2014. Pragmatic borrowing. *Journal of Pragmatics* 67. 17–33.

Andersen, Gisle, Cristiano Furiassi & Biljana Mišić Ilić. 2017. The pragmatic turn in studies of linguistic borrowing. *Journal of Pragmatics* 113. 71–76.

Andersen, Gisle & Anne-Line Graedler. 2020. Morphological borrowing from English to Norwegian: The enigmatic non-possessive-s. *Nordic Journal of Linguistics* 43(1). 3–31.

Androutsopoulos, Jannis. 2012. English 'on top': Discourse functions of English resources in the German mediascape. *Sociolinguistic Studies* 6(2). 209–238.

Arppe, Antti & Juhani Järvikivi. 2007. Every method counts: Combining corpus-based and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3(2). 131–159.

Babel, Anna M. 2016. Affective motivations for borrowing: Performing local identity through loan phonology. *Language & Communication* 49. 70–83.

Backus, Ad. 2020. Usage-based approaches. In Yaron Matras & Evangelia Adamou (eds.), *The Routledge handbook of language contact* (Routledge Handbooks in Linguistics), 110–126. London & New York: Routledge.

Balteiro, Isabel. 2011. A reassessment of traditional lexicographical tools in the light of new corpora: Sports Anglicisms in Spanish. *International Journal of English Studies* 11(2). 23–52.

Balteiro, Isabel. 2014. The influence of English on Spanish fashion terminology: *-ing* forms. *ESP Today: Journal of English for Specific Purposes at Tertiary Level* 2(2). 156–173.

Bontekoning, Aart Cornelis. 2007. *Generaties in organisaties: Een onderzoek naar generatieverschillen en de effecten daarvan op de ontwikkeling van organisaties* [Generations in organizations: A study of generation differences and the effects of it on the development of organizations]. Tilburg: Tilburg University dissertation.

Burmasova, Svetlana. 2010. *Empirische Untersuchung der Anglizismen im Deutschen am Material der Zeitung Die WELT* (Jahrgänge 1994 und 2004). Bamberg: Bamberg University Press.

Cajot, José. 2004. 'Die hollanders toch met hun franse woorden altijd!' Een Vlaams verhaal van de splinter en de balk [Those Dutch with their French words! A Flemish story of the splinter and the beam]. In Johan De Caluwe (ed.), *Schatbewaarder van de Taal. Johan Taeldeman. Liber Amicorum*, 135–153. Ghent: Academia Press.

Campbell-Kibler, Kathryn. 2016. Towards a cognitively realistic model of meaningful sociolinguistic variation. In Anna Babel (ed.), *Awareness and control in sociolinguistic research*, 123–151. Cambridge: Cambridge University Press.

Cheshire, Jenny & Lise-Marie Moser. 1994. English as a cultural symbol: The case of advertisements in French-speaking Switzerland. *Journal of Multilingual and Multicultural Development* 15(6). 451–469.

Clyne, Michael G (ed.). 1992. *Pluricentric languages. Differing norms in different nations*. Berlin & New York: Mouton de Gruyter.

Czech-Rogoyska, Agnieszka & Magdalena Krawiec. 2018. IT-Related Anglicisms in Die Welt and Der Spiegel: A Quantitative Distribution of nouns, Verbs, and Adjectives. *Social Communication* 4(2). 52–58.

De Decker, Benny & Reinhild Vandekerckhove. 2013. De integratie van Engels in Vlaamse jongerentaal kwantitatief en kwalitatief bekeken: Das wel nice! *Nederlandse Taalkunde* 18(1). 2–34.

Devos, Filip. 2018. 'De diehard conservatieve vleugel wordt gesteund door de Tea Party.' Deinvloed van leeftijd op de comprehensie en appreciatie van (het gebruik van) Engelsewoorden in Vlaamse televisiejournaals [The diehard conservative wing is supportedby the Tea Party: The influence of age on the comprehension and appreciation of(the use of) English words in Flemish television news]. In Jacques van Keymeulen, Timothy Colleman, Johan De Caluwe, Veronique De Tier, Anne-Sophie Ghyselen, Roxane Vandenberghe & Ulrike Vogl (eds.), *Woorden om te bewaren: Huldeboek voor Jacques Van Keymeulen*, 643–649. Drongen: Skribis.

Edmonds, Philip & Graeme Hirst. 2002. Near-synonymy and lexical choice. *Computational Linguistics* 28(2). 105–144.

European Commission. 2012. *Special eurobarometer 386: Europeans and their languages*. http://ec.europa.eu/public_opinion/archives/ebs/ebs_386_en.pdf (accessed 27 October 2020).

Garcia-Yeste, Miguel. 2013. The presence and roles of English in Swedish print advertising: An exploratory study. *Nordic Journal of English Studies* 12(1). 65–85.
Gardt, Andreas. 2001. Das Fremde und das Eigene. Versuch einer Systematik des Fremdwortbegriffs in der deutschen Sprachgeschichte. In Gerhard Stickel (ed.), *Neues und Fremdes im deutschen Wortschatz. Aktueller lexikalischer Wandel*, 30–55. Berlin & New York: Mouton de Gruyter.
Garley, Matt. 2010. Realkeepen: Anglicisms in the German hip-hop community. In Marina Terkourafi (ed.), *Languages of global Hip-Hop*, 277–299. New York: Continuum International Publishing Group.
Garrett, Peter. 2010. *Attitudes to language*. Cambridge: Cambridge University Press.
Geeraerts, Dirk & Stefan Grondelaers. 2000. Purism and fashion: French influence on Belgian and Netherlandic Dutch. *Belgian Journal of Linguistics* 13(1). 53–67.
Gerritsen, Marinel, Hubert Korzilius, Frank van Meurs & Inge Gijsbers. 2000. English in Dutch commercials: Not understood and not appreciated. *Journal of Advertising Research* 40(4). 17–31.
Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
González Cruz, María Isabel & María Jesús Rodríguez Medina. 2011. On the pragmatic function of anglicisms in Spanish: A case study. *Revista Alicantina de Estudios Ingleses* 113. 257–273.
Ilić, Biljana Mišić. 2017. Pragmatic borrowing from English into Serbian: Linguistic and sociocultural aspects. *Journal of Pragmatics* 113. 103–115.
Kabatek, Johannes. 1996. *Die Sprecher als Linguisten, interferenz-und Sprachwandelphänomene dargestellt am Galicischen der Gegenwart*. Tübingen: Niemeyer.
Koch, Peter & Wulf Oesterreicher. 2001. Langage parlé et langage écrit. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der Romanistischen Linguistik*, vol. I/2, 584–627. Tübingen: Niemeyer.
Koch, Peter & Wulf Oesterreicher. 2012. Language of immediacy–language of distance: Orality and literacy from the perspective of language theory and linguistic history. In Claudia Lange, Beatrix Weber & Göran Wolf (eds.), *Communicative spaces: Variation, contact, and change: Papers in honour of Ursula Schaefer*, 441–473. Frankfurt: Peter Lang.
Labov, William. 1963. The social motivation of a sound change. *Word* 19. 273–309.
Labov, William. 2001. *Principles of linguistic change*, vol. 2. Social factors. Oxford: Blackwell.
Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner & Samuel Fillenbaum. 1960. Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology* 60(1). 44–51.
Langer, Nils & Agnete Nesse. 2012. Linguistic purism. In Juan M. Hernández-Campoy & Juan Camilo Conde-Silvestre (eds.), *The handbook of historical sociolinguistics*, 607–625. Malden, MA & Oxford: Wiley-Blackwell.
Martin, Elizabeth. 2007. "Frenglish" for sale: Multilingual discourses for addressing today's global consumer. *World Englishes* 26(2). 170–190.
Martin, Elizabeth. 2008. Language-mixing in French print advertising. *Journal of Creative Communications* 3(1). 49–76.
Meurs, Frank van, Berna Hendriks & Dirk Sanders. 2017. Attitudes to English job titles in The Netherlands and Flanders. *Taal en Tongval* 69(1). 89–111.
Meurs, Frank van, Jos Hornikx & Gerben Bossenbroek. 2014. English loanwords and their counterparts in Dutch job advertisements: An experimental study in association overlap. In Eline Zenner & Gitte Kristiansen (eds.), *New perspectives on lexical borrowing:*

Onomasiological, methodological and phraseological innovations, 171–190. Berlin & Boston: De Gruyter Mouton.

Meurs, Frank van, Hubert Korzilius, Brigitte Planken & Steven Fairley. 2007. The effect of English job titles in job advertisements on Dutch respondents. *World Englishes* 26(2). 189–205.

Nejjari, Warda. 2020. *Accents in context. Non-native and native listeners' perceptions and understanding of Dutch-accented English*. Amsterdam: LOT dissertation.

Núñez Nogueroles, Eugenia Esperanza. 2017. Typographical, orthographic and morphological variation of anglicisms in a corpus of Spanish newspaper texts. *Revista Canaria de Estudios Ingleses* 75(noviembre). 175–190.

Onysko, Alexander, Marcus Callies & Eva Ogiermann. 2013. Gender variation of anglicisms in German: The influence of cognitive factors and regional varieties. *Poznan Studies in Contemporary Linguistics* 49(1). 103–136.

Onysko, Alexander & Esme Winter-Froemel. 2011. Necessary loans–luxury loans? Exploring the pragmatic dimension of borrowing. *Journal of Pragmatics* 43(6). 1550–1567.

Pennycook, Alastair. 2003. Global Englishes, rip slyme, and performativity. *Journal of SocioLinguistics* 7(4). 513–533.

Peterson, Elizabeth. 2017. The nativization of pragmatic borrowings in remote language contact situations. *Journal of Pragmatics* 113. 116–126.

Peterson, Elizabeth & Kristy Beers Fägersten (eds.). 2018. Linguistic and pragmatic outcomes of contact with English. *Journal of Pragmatics* 133. 105–108.

Peterson, Elizabeth & Johanna Vaattovaara. 2014. Kiitos and Pliis: The relationship of native and borrowed politeness markers in Finnish. *Journal of Politeness Research* 10(2). 247–269.

Pharao, Nicolai & Tore Kristiansen. 2019. Reflections on the relation between direct/indirect. *Linguistics Vanguard* 5(1). 1–7.

Piller, Ingrid. 2001. Identity constructions in multilingual advertising. *Language in Society* 30(2). 153–186.

Rosseel, Laura, Dirk Speelman & Dirk Geeraerts. 2019. Measuring language attitudes in context: Exploring the potential of the Personalized Implicit Association Test. *Language in Society* 48(3). 429–461.

Rothe, Astrid. 2014. On the variation of gender in nominal language mixings. In Eline Zenner & Gitte Kristiansen (eds.), *New perspectives on lexical borrowing: Onomasiological, methodological and phraseological innovations*, 191–223. Berlin & Boston: De Gruyter Mouton.

Sánchez, José Antonio Jódar & Sirpa Tuomainen. 2014. English in Finland: Opinions andattitudes on the use of anglicisms. *Estudios de Lingüística Inglesa Aplicada (ELIA)*. 97–144.

Sandøy, Helge. 2014. Linguistic globalization: Experiences from the Nordic laboratory. In Eline Zenner & Gitte Kristiansen (eds.), *New perspectives on lexical borrowing: Onomasiological, methodological and phraseological Innovations*, 225–249. Berlin & Boston: De Gruyter Mouton.

Saugera, Valérie. 2017. *Remade in France: Anglicisms in the lexicon and morphology of French*. Oxford: Oxford University Press.

Schaefer, Sarah Josefine. 2019. Anglicisms in German media: Exploring catachrestic and non-catachrestic innovations in radio station imaging. *Lingua* 221. 72–88.

Schmid, Hans-Jörg. 2010. Does frequency in text instantiate entrenchment in the cognitive system? In Dylan Glynn & Kerstin Fischer (eds.), *Quantitative cognitive semantics: Corpus-driven approaches*, 101–135. Berlin & New York: De Gruyter Mouton.

Schmidt, Selina & Stefan Diemer. 2015. Comments on Anglicisms in Spanish and their reception. *Saarland Working Papers in Linguistics* 5(8). 8–16.
Serigos, Jacqueline. 2017. Using distributional semantics in loanword research: A concept-based approach to quantifying semantic specificity of Anglicisms in Spanish. *International Journal of Bilingualism* 21(5). 521–540.
Sijs, Nicoline van der. 1996. *Leenwoordenboe: De invloed van andere talen op het Nederlands* [Dictionary of loanwords: The influence of other languages on Dutch]. The Hague: SdU.
Sijs, Nicoline van der. 1999. *Taaltrots: Purisme in een veertigtal talen* [Language pride: Purisme in forty languages]. Amsterdam & Antwerp: Uitgeverij Contact.
Soares da Silva, Augusto. 2014. Measuring and comparing the use and success of loanwords in Portugal and Brazil: A corpus-based and concept-based sociolectometrical approach. In Eline Zenner & Gitte Kristiansen (eds.), *New perspectives on lexical borrowing: Onomasiological, methodological and phraseological innovations*, 101–141. Berlin & Boston: De Gruyter Mouton.
Spitzmüller, Jürgen. 2007. Staking the claims of identity: Purism, linguistics and the media in post-1990 Germany. *Journal of SocioLinguistics* 11(2). 261–285.
Thøgersen, Jacob. 2004. Attitudes towards the English influx in the Nordic countries: A quantitative investigation. *Nordic Journal of English Studies* 3(2). 23–38.
Thomas, George. 1991. *Linguistic purism*. London: Longman Publishing Group.
Thomason, Sarah. 2001. *Language contact*. Edinburgh: Edinburgh University Press.
Tomasello, Michael. 2005. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Vaattovaara, Johanna & Elizabeth Peterson. 2019. Same old paska or new shit? On the stylistic boundaries and social meaning potentials of swearing loanwords in Finnish. *Ampersand* 6. 1–9.
Vandenbussche, Wim, Willemyns Roland, Jetje De Groof & Eline Vanhecke. 2005. Taming thistles and weed amidst the wheat: Language gardening in 19th century Flanders. In Nils Langer & Winifred V. Davies (eds.), *Linguistic purism in the Germanic languages*, 46–61. Berlin & New York: Walter de Gruyter.
Verhagen, Véronique. 2019. *Illuminating variation. Individual differences in entrenchment of multi-word units*. Amsterdam: LOT dissertation.
Walsh, Olivia. 2014. 'Les anglicismes polluent la langue française': Purist attitudes in France and Quebec. *Journal of French Language Studies* 24(3). 423–449.
Walsh, Olivia. 2015. Attitudes towards English in France. In Andrew Linn, Neil Bermel & Gibson Ferguson (eds.), *Attitudes towards English in Europe*, 27–54. Berlin & Boston: De Gruyter Mouton.
Winter, Bodo. 2019. *Statistics for linguists: An introduction using R*. London & New York: Routledge.
Winter-Froemel, Esme. 2008. Unpleasant, unnecessary, unintelligible? Cognitive and communicative criteria for evaluating borrowings and alternative strategies. In Roswitha Fischer & Hanna Pułaczewska (eds.), *Anglicisms in Europe: Linguistic diversity in a global context*, 16–41. Cambridge: Cambridge Scholars Publishing.
Winter-Froemel, Esme. 2011. *Entlehnung in der Kommunikation und im Sprachwandel. Theorie und Analysen zum Französischen*. Berlin & Boston: De Gruyter.
Winter-Froemel, Esme. 2017. The pragmatic necessity of borrowing: Euphemism, dysphemism, playfulness – and naming. *Taal en Tongval* 69(1). 17–46.

Winter-Froemel, Esme. 2018. Reanalysis in language contact: Perceptive ambiguity, salience, and catachrestic reinterpretation. In Eline Zenner, Ad Backus & Esme Winter-Froemel (eds.), *Cognitive contact linguistics: Placing usage, meaning and mind at the core of contact-induced variation and change*, vol. 62, 81–126. Berlin & Boston: De Gruyter Mouton.

Winter-Froemel, Esme. 2020. Les néologismes, un sous-type de l'innovation lexicale: Réflexions à partir d'études de corpus généraux dans trois langues romanes. *Neologica* 14. 25–46.

Winter-Froemel, Esme, Alexander Onysko & Andreea Calude. 2014. Why some non-catachrestic borrowings are more successful than others: A case study of English loans in German. In Amei Koll-Stobbe & Sebastian Knospe (eds.), *Language contact around the globe*, 119–142. Frankfurt: Peter Lang.

Wolf, Hans-Georg & Frank Polzenhagen. 2009. *World englishes: A cognitive sociolinguistic approach*. Berlin & Boston: De Gruyter. Mouton.

Yang, Wenliang. 1990. *Anglizismen im Deutschen*. Tübingen: Max Niemeyer Verlag.

Zenner, Eline & Gitte Kristiansen. 2014. *New perspectives on Lexical borrowing: Onomasiological, methodological and phraseological innovations*. Berlin & Boston: De Gruyter Mouton.

Zenner, Eline, Laura Rosseel & Andreea Simona Calude. 2019. The social meaning potential of loanwords: Empirical explorations of lexical borrowing as expression of (social) identity. *Ampersand* 6. 1–14.

Zenner, Eline, Laura Rosseel & Dirk Speelman. 2020. Starman or Sterrenman: An acquisitional perspective on the social meaning of English in Flanders. *International Journal of Bilingualism* 25(3). 568–591.

Zenner, Eline, Dirk Speelman & Dirk Geeraerts. 2012. Cognitive sociolinguistics meets loanword research: Measuring variation in the success of anglicisms in Dutch. *Cognitive Linguistics* 23(4). 749–792.

Zenner, Eline, Dirk Speelman & Dirk Geeraerts. 2013. Macro and micro perspectives on the distribution of English in Dutch: A quantitative usage-based analysis of job ads. *Linguistics* 51(5). 1019–1064.

Zenner, Eline, Dirk Speelman & Dirk Geeraerts. 2015. A sociolinguistic analysis of borrowing in weak contact situations: English loanwords and phrases in expressive utterances in a Dutch reality TV show. *International Journal of Bilingualism* 19(3). 333–346.

Supplementary Material: The data supporting this study may be found at https://doi.org/10.5281/zenodo.4815367.