Chapter 5

Namibian German and gender: A corpus study on the use of transferred lexical items

Yannic Bracke

Friedrich-Schiller-Universität Jena

This chapter presents a quantitative corpus study of informal speech from male and female adolescent and adult Namibians with L1 German. A key feature of Namibian German is various forms of language mixing, mostly with material from English and Afrikaans. Previous sociolinguistic research, as well as statements by community members, suggest that male speakers might use more other-language material in their speech. I identified other-language material in a corpus of peer group conversations by Namibian German adolescents and adults and investigated the amount of transferred lexical items (other-language material excluding multi-word code-switches) that speakers of different age and gender used. Furthermore, I analyzed the proportion of the donor languages English and Afrikaans. Concerning the frequency of transferred lexical items, the results show an age difference between younger and older speakers, but fewer clear differences between speakers of different gender. English is the prime donor language in all groups, but subtle differences in the proportion of Afrikaans may point to interesting sociolinguistic dynamics.

1 Introduction

Namibian German (NG), or Namdeutsch, has been gaining interest as a linguistic research topic in recent years. The fact that German is spoken in Namibia mainly goes back to the immigration of German-speaking people to the area of present-day Namibia during and after its time as a colony of the German Reich

between 1884 and 1915. Today the descendants of these immigrants form a minority of approximately 20,000 L1 speakers of German. They use the language in a variety of contexts, private and official, for example in schools, clubs and churches. That said, they live in a highly multilingual country and typically also speak English and Afrikaans. Today, English is the country's sole official language while Afrikaans was an official language before Namibia's independence from the South African apartheid regime in 1990, and continues to be used as a lingua franca. Some members of the NG community also speak Bantu and Khoisan languages, such as Oshiwambo, Herero, or Nama/Damara, but this is markedly less common (see Shah & Zappen-Thomson 2018 and Zimmer 2019 for more detailed descriptions of the general situation of German in Namibia). One effect of this multilingual situation is the occurrence of various forms of language mixing in Namibian German. In particular, English and Afrikaans influence NG language use (see, e.g., Shah 2007: 22; Wiese et al. 2014; Wiese et al. 2017; Zimmer 2019: 1185–1187). Three examples are provided below.[1]

(1) es   hilft   nich nur mein **self esteem i love doing it**
     it.NOM help.3SG not only my   **self esteem I love doing it**
     'Not only does it help my self esteem, I love doing it.' [NAM119W1]

(2) ach   dis is   n   **gesükkel**
     EXCLAM this be.3SG INDEF.SG.NEUT.NOM **struggle**
     'Oh, this is exhausting.' [NAM155M2]

(3) couscous is   **mooi**
     couscous be.3SG.PRES **good**
     'Couscous is tasty.' [NAM006M1]

As a key feature of NG, language mixing phenomena have already received a substantial amount of scholarly attention. It has been argued that individual words from English or Afrikaans have gained the status of accepted loanwords and are also used in formal registers (Kellermeier-Rehbein 2016: 225–226). In general, however, the use of other-language material appears to be mostly reserved

---

[1]All examples in this chapter are taken from the corpus *Deutsch in Namibia* (see below). The German original is provided in cGAT transcription (Schmidt et al. 2015), followed by an interlinear gloss and an English translation in natural language. For the sake of anonymity speaker names have been replaced by an alias (next to the translation) that provides some information. Aliases are prefixed with *NAM*, followed by three digits for identification. The following letter denotes the speaker gender (*M* for male, *W* for female) and the final digit denotes one of four age groups (*1*: 20 years or younger; *2*: 21–40 years; *3*: 41–60 years, *4*: 61 years or older).

for informal settings (Wiese & Bracke 2021). This is connected to community attitudes towards mixing and linguistic purism: Standard German as it is (imagined to be) spoken in Germany is regarded as the prestige variety by German-speaking Namibians (Zimmer 2019: 1185). Other-language influences are often stigmatized as markers of substandard and "bad German" (see Gregersen & Langer 2021 [this volume] for a more detailed account of linguistic purism in language contact situations). At the same time, these influences also bear some positive connotations for many speakers who associate them with a specific Namibian German identity, setting Namibian Germans apart from Germans in Germany (Schmidt-Lauber 1998: 308–309; Wiese & Bracke 2021; Radke 2021 [this volume]).

Various recent studies have investigated language mixing phenomena in NG with some pragmatic or sociolinguistic focus, analyzing them in the context of register variation (Wiese & Bracke 2021), in online media and communication (Radke 2017; 2021 [this volume]), in youth language (Kellermeier-Rehbein 2015; 2016), and with regard to speaker age (Zimmer forthcoming). This line of work is extended in this chapter by focusing on speaker gender, an aspect that has so far been largely neglected for NG. Specifically, I present a quantitative corpus study on the use of transferred lexical items by male and female speakers in informal peer group conversations.

What do I mean by *transferred lexical items* in this chapter? The literature on language contact phenomena provides a variety of terms and concepts but it does not agree on their exact definition in all cases. There is no controversy about the concept of established loanwords. These are words originating from a different donor language, which have become part of a recipient language's lexicon, which are integrated into its grammatical system, and which are also used by monolingual speakers in the community.[2] It is equally agreed that the use of established loanwords differs from code-switching (CS), which is often defined as the "juxtaposition" of two or more languages (Poplack 2004: 589; Auer 2011: 460). Example (1) marks an unequivocal case of CS as it contains a stretch of English words showing no integration and retaining English grammar. By contrast, the concept of borrowing and its demarcation from CS causes some controversy. Poplack and colleagues have argued that, aside from the borrowing of established loanwords, words can also be borrowed "for the nonce" (Poplack et al. 1988; Poplack 2004; 2018). These "nonce borrowings" are similar to established loans in some ways and similar to CS in other ways:

---

[2]As examples of established loanwords in NG, Zimmer (2019: 1185) lists *Rivier* ('dry river') and *braaien* ('to barbecue'), both from Afrikaans. Note, however, that in the case of NG the loanword criterion referring to monolingual community members can be largely disregarded because almost all community members are multilingual (see above).

> Like its established counterpart, the nonce borrowing tends to involve lone lexical items, generally major-class content words, and to assume the morphological, syntactic, and optionally, phonological identity of the recipient language. Like CS, on the other hand, particular nonce borrowings are neither recurrent nor widespread, and nonce borrowing necessarily requires a certain level of bilingual competence. (Poplack 2004: 590)

In this view, being "neither recurrent nor widespread", *Gesükkel* in example (2) constitutes a nonce borrowing, since it is a derivation of the Afrikaans verb *sukkel* ('to struggle') to a noun via a German prefix. However, even a word that is not overtly integrated, like the Afrikaans *mooi* ('good') in example (3), could be a nonce borrowing because a native German adjective would not bear overt morphological inflection in the same slot. It would be a nonce borrowing *candidate* because the possibility of single-word CS is not excluded a priori for words that are not overtly integrated in this view (Poplack et al. 1988: 7). However, other researchers reject the concept of nonce borrowings, contending that every use of lexical material from another language that is not an established loanword can be subsumed under the concept CS. Specifically, they see single-word or even single-morpheme other-language material simply as a very short form of CS (Myers-Scotton 2002: 154–7; Haspelmath 2009: 41). I follow the intuition of Poplack and others that nonce borrowings should be distinguished from CS.[3] Yet, even if all single unintegrated other-language items were referred to as CS, they are arguably different from multi-word CS in some ways. In terms of psycholinguistic activation, a longer sequence likely leads to a stronger activation of the other language, while German would stay most activated during short sequences (Muysken 2000: 8, 34). From a sociolinguistic perspective, the longer the other-language stretch lasts, the more the multilingualism is foregrounded in the conversation. Thus, in order to narrow down the range of phenomena that are analyzed in the study, I decided to exclude multi-word CS from the analysis. Consequently, the data that is analyzed includes established loanwords as well as (candidates for) nonce borrowing and/or single word CS, depending on the point of view.[4] For the purposes of this chapter, I subsume this material under the term *transferred lexical items*. The material was identified with the help of an annotation system (see §4.2), which facilitated the operationalization of transferred lexical items (see §4.3).

---

[3]A related distinction is the one between *insertion* and *alternation* made by Muysken (2000).

[4]Note that in this chapter I do not determine to which of the categories each token belongs. This would be the task of a separate article. The purpose of the study presented here was to focus on the sociolinguistic variation concerning the use of the analyzed material.

The study is based on analyses of some 100,000 tokens from the corpus *Deutsch in Namibia* (*DNam*), a recent collection of spoken data by German-speaking Namibians (Zimmer et al. 2020). Since some parts of this corpus only contain data by adolescent speakers (aged 14 to 18) and others only contain data by adult speakers (aged 26 to 65), it was possible to examine the role of gender in both of these (broad) age groups separately and compare the results.

The study's main focus is on the frequency of transferred tokens in speech. Due to previous work and my own observations concerning the community, which are laid out in the next section, my goal was to examine whether male speakers use more transferred lexical items than female speakers. Concerning the aspect of age, I expected that adolescents used more transferred lexical items than adults, because younger speakers are often linguistically more creative (Wiese et al. 2014: 277) and because it has been suggested by researchers and community members that Namibian German youth language is particularly rich with influences from other languages (Kellermeier-Rehbein 2016: 228). Despite this, a previous study on loanwords in NG translations of "Wenker sentences" does *not* report that younger speakers use more loanwords than older speakers (Zimmer forthcoming).[5] An additional sociolinguistic variable taken into account in the present study is the type of school that the adolescent speakers attended. The role of German and the amount of German language instruction varies substantially between different schools in Namibia, which might be reflected in the use of transferred material.

The analysis of the frequency of transferred lexical items in general is complemented by an additional investigation of the proportion of transferred lexical items from different donor languages. Previously, it has been reported that Afrikaans is NG's most important donor language, ranking above English, and that both Afrikaans and English rank above Bantu and Khoisan languages (Nöckler 1963: 47; Böhm 2003: 568; Kellermeier-Rehbein 2016: 229–230). However, it has been assumed that English may be in the process of overtaking Afrikaans because, unlike Afrikaans, it is not associated with apartheid and receives more institutional support in post-independence Namibia (cf. Shah 2007: 43; Kellermeier-Rehbein 2016: 230; Zimmer forthcoming). This corresponds to reports by community members claiming that younger speakers are more influenced by English than older speakers (cf. Zimmer forthcoming). Yet, in the quantitative "Wenker sentences" study no such tendency is observed and Afrikaans seems far ahead

---

[5]Instead, Zimmer (forthcoming) reports a U-shaped pattern of age differentiation with middle-aged speakers (40–49 years) using markedly fewer loanwords than younger and older informants.

of English. The proportion of Afrikaans vs. English tokens is approximately 80 to 20 percent in all age groups (Zimmer forthcoming). To my knowledge, there were no previous indications of any gender-specific differences with respect to the proportion of donor languages.

The chapter is structured as follows. In the following section I address why I think it is worth analyzing the use of transferred lexical items in NG with respect to speaker gender. Then, a general description of the data is given (§3) and the methodology is presented (§4). In the main section the results of the corpus study are presented (§5). I conclude the chapter with a summary and discussion of the results and some perspectives for future research (§6).

2 Why gender?

The relationship between gender and language use has been a subject of soci-olinguistic research for several decades (cf. Coates & Pichler 2011). Today, it is widely accepted that gender is a socially constructed category, but one that plays a powerful role in people's lives nonetheless. As such, it is also meaningful for how people speak and why they speak the way they speak. It is clear, however, that gender is not a category that exists independent of other social categories or circumstances (Eckert 2011: 65) and I did not assume that gender is the sole or primary variable influencing the use of transferred material in NG. This is why I looked at gender in conjunction with age, and point to other sociolinguistic cir-cumstances in the analysis. Nonetheless, there are certain aspects that indicate a possible existence of gender differences when it comes to the use of transferred lexical items in NG.

I have said above that language mixing phenomena are viewed as markers of nonstandard speech in the NG community. Early quantitative sociolinguistic studies from Britain and the US related nonstandard speech to gender by report-ing almost unanimously that low prestige, nonstandard variants are favored by male rather than by female speakers (Wolfram 1969; Trudgill 1972; Macaulay 1977; Labov 2006). Hence, analyzing the use of transferred lexical items in the speech of male and female speakers presents an opportunity to investigate whether or not this commonly observed pattern is also found in the NG speech com-munity. A first quantitative piece of evidence comes from Zimmer's (forthcom-ing) "Wenker sentences" study where he reports that 20–29- and 30–39-year-old women used markedly fewer loanwords in their translations than men of the same age.[6] Some studies conducted in other communities have found that male

---

[6]The author reasons that these differences emerge from a conservative distribution of roles in childcare and that women in these age groups use fewer nonstandard words because of their role as mothers (Zimmer forthcoming).

speakers already tend towards more nonstandard speech in adolescence (Eckert 2003: 387; Eisikovitz 2011) underlining the importance and value of differentiated analyses. For NG there is no primary evidence of this, except for a few statements in interviews with NG adolescents that Heike Wiese conducted in 2013.[7] Some interviewees reported that boys used more "slang", while others denied that linguistic differences between boys and girls existed.[8]

Support for the assumption that male NG speakers might use more transferred lexical items than female speakers comes from narratives by community members. In an extensive ethnography on the Namibian German community Schmidt-Lauber (1998) addresses the role of gender and gender(ed) stereotypes. She reports that, in her interviews with community members, it became apparent that "the Farmer" serves as an ideal and a figure of identification for many members of the group. This stereotypical figure is characterized by its toughness and masculinity (Schmidt-Lauber 1998: 236–240). Similar sentiments can be found in an excerpt from an interview in the *DNam* corpus where the interviewee, a 31-year-old man, talks about the terms *braaien* (from Afrikaans *braai*, 'barbecue') and the Standard German translation *grillen* (cf. also Zimmer forthcoming). Community members widely regard *braaien* as an important practice that seems to establish a sense of community and identity (cf. statements in the *DNam* interview corpus).[9]

> Also, wenn […] ein Freund von mir jetzt da mit "grillen" anfängt so, da würde ich sagen: "Was ist mit dir los?" […] Weil "braaien" ist einfach […] so Tradition hier […]. Wir machen das […] auch nicht mit Kohle sondern mit Holz und richtig schön Fleisch. Und nicht so nach […] deutschem Gesetz, sag ich mal, jetzt bloß kein Rauch und bloß nicht dies und das und jenes. Bei uns muss das richtig son *Männerding* sein. Und wenn jetzt einer kommt mit 'grillen', dann hört sich das wieder so […] *verweiblicht* an, sag ich jetzt mal.

> 'Well, if a friend of mine said "grillen", I would say "What's the matter with you?" Because "braaien" is just a tradition here. We don't use charcoal but wood and we use proper meat. We don't stick to "German Law", so to speak, making sure there is no smoke and no this and no that. Here it has to be a real *guy thing*. And if somebody comes along and calls it "grillen" then this sounds somehow *effeminate*, so to speak.'

---

[7]Other aspects of these interviews are mentioned in Wiese et al. (2017).

[8]Also, Deumert (2009: 359) reports in passing that male students make stronger use of Namibian German variants, but she does not provide quantitative evidence for this claim.

[9]Since the content of this quote is more important than its form, I provide a reader-friendly normalized transcription where hesitations, etc. are left out (marked by *[…]* in the German original) and punctuation and emphasis is added.

In his statement the interviewee links Namibian German identity, masculinity and language. Firstly, he marks *braaien* as a distinctly Namibian German practice by setting Namibian Germans apart form Germans from Germany, characterizing the former as tough – by virtue of using wood, fire and "proper meat" – and the latter as law-abiding. Secondly, the loan *braaien* is presented as the only proper way to refer to this practice. Thirdly, the standard German term *grillen* is disdained as "effeminate" and *braaien* is called a "guy thing", thereby linking practice and terminology to masculinity. Here, the apparent ideological connection between a traditional Namibian German identity and traditional ideas of masculinity includes the use of NG-specific lexis. As in the case of *braaien*, NG-specific lexis is typically transferred lexis.[10] This raises the question whether such a connection between masculinity and transferred lexis is more widespread and if it is reflected in everyday language use. Do male speakers use more transferred lexical items in their speech?[11]

It is important to note that more recent approaches to language and gender regard linguistic variation not as a mere reflection of social variables, but as one of the resources that people use to construct their identity (cf. contributions in Eckert & McConnell-Ginet 1999; Coates & Pichler 2011; Eckert 2012). This assumption is particularly plausible in the context of language mixing phenomena of which speakers are largely conscious. That is, potential linguistic differences between male and female speakers could be interpreted to be part of the practice of *doing gender*.

The *DNam* corpus allows us to investigate whether differences exist and, if so, where. Analyzing language and gender in the Namibian German context should not stop with this corpus linguistic investigation. For a thorough understanding of different behavior it is necessary to engage in detailed qualitative analyses (cf. Eckert & McConnell-Ginet 1999; Eckert 2012). However, the quantitative approach I take here is a first step towards an understanding of the role of gender in Namibian German. In what follows the study is presented in detail, starting with a description of the data that were used.

---

[10] A related observation concerns the term *oukie*. The common meaning of the word is 'guy' (from Afrikaans *ou*) referring to men only. Yet, *oukies* may also be used as a seemingly generic term for members of the NG community regardless of gender (Pütz 1982; see Radke 2021 [this volume] for a detailed discussion of the terms usage), suggesting that the prototypical Namibian German is male. This is underlined by attestations like "oukies and ladies" to refer to Namibian Germans of both gender (Radke 2021 [this volume]), since this suggests that there is a need to use an additional term (*ladies*) in order to explicitly refer to female Namibian Germans as well.

[11] Similarly, the so called *covert prestige hypothesis*, which Trudgill (1972) developed as a possible explanation for the gender differences observed in the Anglo-American studies cited above, assumes that gender-specific associations of nonstandard speech might have contributed to the observed tendencies (Trudgill 1972: 183).

3 Data

The data for this study came from a pre-release version of the corpus *Deutsch in Namibia* (*DNam*, 'German in Namibia'). The released corpus is accessible online via the *Datenbank für Gesprochenes Deutsch* (Zimmer et al. 2020).[12] *DNam* is a systematic collection of the language use of the German-speaking minority in Namibia. The corpus is complemented by a collection of sociolinguistic metadata about the speakers, which is crucial for investigations like the one presented in this chapter. *DNam* contains audio files and transcriptions of three kinds: sociolinguistic interviews, an experimental set-up ("language situations", cf. Wiese 2020; Wiese & Bracke 2021), and free peer group conversations that were recorded in the absence of researchers. The latter were chosen as a corpus for this study because this part of the *DNam* corpus is arguably the closest approximation to everyday linguistic interactions of Namibian Germans available for research. A minor drawback of the free conversations in the context of this study is that the content of the conversation can influence the likelihood of (specific) transferred content words. For example, in German conversations about technical equipment the usage of English terms is more likely than in other contexts. The possibility of this effect has to be kept in mind, however I regarded the use of authentic language as a priority. Moreover, the fact that topics usually changed during conversations and that the corpus consists of multiple conversations has presumably mitigated this effect.

The free conversations consist of two subcorpora, one that contains only conversations of adolescents and one that contains only conversations of adults. This allowed for a separate investigation of gender differences in peer group conversations of both groups. The adolescent subcorpus consists of 13 conversations by groups of three to five people, while the adult subcorpus consists of six conversations by groups of two or three people. Consequently, the number of contributing speakers is substantially larger in the adolescent subcorpus. This is displayed in Table 1, together with information about the gender and age of speakers in both subcorpora.[13] It should be noted that the number of male and female speakers is not perfectly balanced in either subcorporus. The conversations of adults were recorded on farms and in the cities of Windhoek, Otjiwarongo and Omaruru. The adolescents were recorded in schools in Windhoek, Otjiwarongo and Swakopmund with the exception of six speakers, who were recorded during private German lessons in Windhoek (see §5.1).

---

[12]See https://dgd.ids-mannheim.de (11 May, 2021).
[13]For two adolescents the exact age is not documented.

Table 1: Speaker gender and age in the subcorpora

	Total	Female	Male	Age span	Mean age (std)
Adolescents	51	32	19	14−18*	16.1* (1.2)
Adults	14	8	6	26−65	42.2 (13.7)

The conversations of adolescents have a mean length of approximately 36 minutes and a total length of approximately seven hours and 43 minutes. The conversations of adults are shorter with a mean length of approximately 13 minutes, resulting in a total length of one hour and 20 minutes. The audio data has been transcribed according to the cGAT transcription system (Schmidt et al. 2015). The token and type counts in both subcorpora are displayed in Table 2.

Table 2: Token and type count of the two corpora

	Tokens	Types
Adolescents	90998	7350
Adults	14672	2074
Total	105670	8189

4 Methodology

4.1 Approach

The general approach of this study is that of a quantitative corpus analysis. Specifically, I calculated token frequencies for transferred lexical items and other tokens and used statistical methods to explore the assumptions concerning age and gender discussed above. I deal with details of the identification of other-language material in the corpus and of the operationalization of transferred lexical items in the two next sections.

Note that this study is not variationist in the sense of the studies by Labov and others, mentioned above, because I did not look at a single variable with a few variants, such as the realizations of a phoneme. Rather, I investigated how common the use of transferred lexical items in general is in the speech of NG speakers.[14]

---

[14]Technically, it would be possible to apply the variationist *principle of accountability* (e.g., Tagliamonte 2012: 9) to such an analysis. This would mean that every occurrence of every

I took two different statistical approaches to the corpus data. In a classical corpus linguistic approach every token (or type) in the corpus is treated as an observation having a binary property – being a transferred lexical item (as operationalized below) or not – and having been uttered by a member of some group (females, adolescents, etc.), with the group being treated as a subcorpus. The result is a neat two-by-two frequency table and a relative frequency of transferred tokens (per 1000 tokens) in the entire group subcorpus. Differences between groups can be tested for significance with a χ^2 test and effect sizes can be evaluated with the correlation coefficient Cramer's φ.

The pitfall of the corpus linguistic approach is that group-internal differences cannot be evaluated since the behavior of individual speakers is not taken into account. However, the individuals in each group may have contributed to the total values in very different ways. In the most extreme case all transferred tokens could have been uttered by a single individual while the rest used none – a fact that would go unnoticed if only this approach was used. Quantitative studies on language and gender have been criticized for not analyzing intra-group differences and foregrounding inter-group differences (Eckert & McConnell-Ginet 1999: 193; cf. also Baker 2014 for a critique that specifically targets corpus linguistic approaches). For this reason, I employed an additional approach. Here, the relative amount of transferred tokens (per 1000 tokens) for each individual speaker was calculated. The resulting values for each person can be compared to members of their own group (gender, age, school) as well as to members of the other group. The non-parametric Wilcoxon rank-sum test was used for testing whether significant differences exist between sets of individual values. In the results section, I show how this approach constituted a valuable complement to the first one.

A final point to note concerning the statistical analysis, is that the group of speakers in the corpus is neither homogeneous nor balanced with respect to their sociolinguistic background. For example, the adults have different occupations and the students go to different types of schools. Moreover, the number of male and female participants for each occupation/school type varies. In the analysis, I show that this imbalance must be taken into account before drawing premature conclusions from the data.

---

other-language type in the corpus would have to be measured together with all potential Standard German alternatives. In some cases, such as the discourse marker *like* (Tagliamonte 2012: 270) that appears in NG as well (Wiese et al. 2014: 299–300), an item can also be substituted by a null-variant, meaning that in these cases, every single possible place of occurrence would have to be identified. Needless to say, such an approach would have exceeded the capacities available for this study by far.

4.2 Annotation

A nontrivial prerequisite for conducting a corpus study on the use of transferred lexical items is the identification of tokens of interest. For this purpose, I made use of an annotation system, which I present in this section. Due to the terminological and conceptual controversies over the categorization of other-language material mentioned in the introduction, the annotation system was deliberately designed to be theory-neutral, refraining from the use of the terms *code-switching* and *borrowing*. It is based on directly observable criteria as much as possible. Still, the annotations can be used to approximate these concepts to some extent as I explain in the next section. The annotations were eventually added to all parts of the *DNam* corpus and will be made available to the public in a future version of the corpus. This will allow researchers to perform studies similar to this one, potentially choosing their own selection of other-language tokens.

The most fundamental feature of the annotation is the identification of other-language tokens and a secondary feature is assigning different properties to them. An other-language token was defined in the broadest possible sense as any token that can be identified as originating from a Namibian language other than German, such as English, Afrikaans, Oshiwambo, and so on.[15] For the purpose of distinguishing different kinds of other-language tokens the annotation-tag for each token encoded the following four properties:

1. DONOR LANGUAGE: From which language does the token originate?

2. SEQUENCE (±): Is the token part of a sequence of more than one other-language token (from the same donor language)?

3. INTEGRATION (±): Is the token overtly integrated into German on a morphological level or not?

4. DICTIONARY ENTRY (±): Does an entry for the token's lemma exist in a monolingual dictionary for Standard German (*Duden*) or not?

In order to be assigned the property +SEQUENCE, a token needed to have at least one direct neighbor token from the same donor language. Between direct neighbors there are no speaker transitions, no distinguishable (i.e., transcribed) pauses and no tokens from any other language. Importantly, a compound lexeme consisting of two separate graphematic words (e.g., *phone cover*) was treated as

---

[15]This criterion and some other aspects of the annotation system were inspired by a workflow described in Poplack (2018: 42).

a single token and was annotated as –SEQUENCE unless the above characteristics applied also to the compound. The property +INTEGRATION was assigned either if a token showed overt German inflection that would not occur in the donor language (e.g., *mooi-es stück*, 'good piece') or German derivation (e.g., *Gesükkel*, 'struggle'), or if it was a compound of a native German and an other-language lexeme (e.g., *Babystimme*, 'baby voice'). Tokens from non-inflectible word classes have the property –INTEGRATION (not integrated) by definition, as they can never be overtly integrated. For the property DICTIONARY ENTRY I employed the online version of the *Duden*.[16] A Namibian German codex does not exist and, as stated above, standard language in Germany (which the *Duden* aims to document) serves as a linguistic point of reference in the NG community.[17] In case an other-language token was used as part of a constructed dialogue (i.e., a report of real or fictional speech) or in a meta-linguistic way, this was encoded in the annotation-tag as well so that these cases could be excluded as they diverge from "normal" language use.[18]

For reasons of space, I will not go into a more detailed description of the annotation process. In the next section I explain how the annotation system was used to select the relevant data for analysis.

4.3 Operationalizing transferred lexical items

The annotations described in the preceding section identified every token in the free conversations corpus originating from a local language other than German. The total number of such other-language tokens in the corpus is 6906. I stated in the introduction that these tokens constitute several different language mixing phenomena and have motivated my decision to exclude instances of multi-word CS from the data I analyzed. The annotations allowed me to define multi-word CS for the purposes of this chapter as follows: Any sequence of two or more words from one donor language (+SEQUENCE) in which all tokens have the property –INTEGRATION and at least one token does not have a dictionary entry (–DICTIONARY ENTRY). This definition was adopted here because I regard inte-

---

[16]The annotation process took place between January and April 2019. That is, a word had to have an entry in the online *Duden* (https://www.duden.de) during that time in order to be annotated as +DICTIONARY ENTRY.

[17]The status of the *Duden* in the Namibian German community is exemplified by the fact that it is featured in the school logo of the German private school in Omaruru (cf. http://www.dpso.iway.na/index.html, 16 June, 2021).

[18]650 tokens were excluded from the analysis because of this. They are already excluded from the counts presented in Table 2.

gration as indicative of (nonce) borrowing and a dictionary entry as indicative of an established loanword.[19]

The remaining other-language tokens that are not understood as instances of multi-word CS are referred to as *transferred lexical items* (or simply *transferred tokens/words*) and constitute the data that is analyzed in the following sections (5036 tokens). Specifically, as I indicated in the introduction, this means that any single lexical item (−SEQUENCE) is included. Single other-language words can be morphologically integrated, as in (2), or dictionary-attested, as in (6), or both, as in (7), or neither, as in (3). Note that all dictionary-attested words were included, without discriminating between long established loanwords (like *okay*) and recently established ones (like *chillen*). Despite their attestation, these words are ultimately of non-German origin and speakers are often aware of this.[20] Note further that while dictionary attestation is interpreted as a sign of an established loanword here, it is not a necessary condition. In particular, Afrikaans-origin words can have the status of established loanwords (as has been noted in the literature, see §1), but there are virtually no dictionary-entries for Afrikaans-origin words in the *Duden* so that the method adopted here has no way of identifying established loans from Afrikaans a priori.

The single other-language items constitute the vast majority (96.8%) of the transferred lexical items in the corpus. The rest are tokens from sequences of two or more words for which the two other criteria for multi-word CS do not hold. Thus, these are tokens from sequences that consist of one or more integrated items, as in (4), or that are made-up entirely of dictionary-attested words, as in (5).[21]

(4) weil   wir   sonst   **absent ge-marked**   werdn
     because we.NOM otherwise **absent PTCP-marked** be.3PL.PRES
     'Because otherwise we will be marked as absent.' [NAM118W1]

(5) achso   so da/ **okay okay cool**
     EXCLAM SO (?) **okay okay cool**
     'Oh okay okay cool.' [NAM171W2]

---

[19]Other definitions of (multi-word) CS are conceivable, borderline cases may be excluded by this definition. However, its advantage lies in its simplicity and transparency.

[20]Since all speakers are fluent in English, it can only remain speculative whether the English or the German lexicon was accessed for the use of such a word in a particular instance (cf. Haspelmath 2009: 40). Moreover, the decision not to exclude specific words, however common (e.g., *okay*), from the data selection ensured maximal transparency, as there was no need to establish a (necessarily subjective) definition of commonness.

[21]Because of these features, these sequences could be regarded as a series of borrowings.

(6) ich   war   der   **boss**
     I.NOM be.1SG.PRET DEF.SG.MASC.NOM **boss**
     'I was the boss.' [NAM016W1]

(7) ich   hab   viel   zu wenig mit   dem   **ge-chill-t**
     I.NOM have.1SG.PRES much too little   with DEM.SG.MASC **PTCP-chill-PTCP**
     'I have chilled with him far too rarely.' [NAM173M2]

Some similar studies have excluded other-language tokens from their data set for which no proper native alternative existed (cf. Zenner et al. 2015: 336; Calude et al. 2017). However, this must inevitably lead to problems when trying to define this criterion: What should count as a proper equivalent of a German word and what not?[22] The problem is that, presumably, there are almost no cases where an other-language word does not have at least a slightly different meaning, be it denotative or connotative, than its German counterpart (cf. Zimmer 2019: 1185). Accordingly, I refrained from this practice, keeping the definition of the analyzed material simple and transparent. In the following section I present the results of the analyses conducted on this material.

5 Results

5.1 Frequency of transferred lexical items

The first part of this section is concerned with the frequency of transferred tokens in the speech of Namibian Germans in the *DNam* corpus. As predicted, the results show that adolescent speakers used more transferred lexical items than adult speakers. This is illustrated in Table 3. The relative frequency of transferred tokens (recall that instances of multi-word CS are excluded from this category) per 1000 tokens, displayed in square brackets, is higher by almost 20 tokens in the adolescents subcorpus. The absolute frequencies show the distribution of all tokens over the categories *transferred* and *other*. The difference in this distribution is statistically significant ($\chi^2 = 104.874$; $p < 0.001$).[23] The effect size φ of 0.032 seems to indicate an extremely marginal effect since its potential values range between 0 and 1. Yet, it should be noted that in a corpus of conversations that are (predominantly) in German, the number of native German tokens is of course much higher than the number of transferred tokens. In this respect both samples would always resemble each other to a high degree. To illustrate this, consider

---

[22]Cf. considerations in (the second part of) Fn. 20.

[23]For all χ^2 test results reported in this chapter the degrees of freedom (df) equal 1.

the following hypothetical situation: If the adults had used only five transferred tokens in total, and the adolescents had used the same amount as they actually did, the test statistics would be extremely high ($\chi^2 > 761$) while the effect size would remain below 0.1 ($\varphi = 0.085$). This has to be kept in mind for other effect sizes that I report in this section as well.

Table 3: Distribution of transferred and other tokens over adolescents and adults. [rel.] = relative freq. per 1000 tokens

	Tokens	
	Transferred [rel.]	Other
Adolescents	4582 [50.4]	86416
Adults	454 [30.9]	14218

A look at the frequencies of the individuals in each age group generally confirms the results presented above, but adds further insight. It shows, for example, that all individuals used some amount of transferred tokens, although for a few individuals in both age groups this amount is very small (less than seven transferred tokens per 1000 tokens). Generally most adolescents used more transferred lexical items than the adults in the sample (see Figure 1). According to a one-tailed Wilcoxon rank-sum test, this is a significant difference ($W = 466$; $p = 0.041$).[24] The individual frequencies also reveal that the three youngest adults have the highest relative frequency of transferred words. These are two 26-year-old men and a 31-year-old woman from Windhoek who took part in the same conversation together. Their use of transferred words is similar to, or even higher than, that of many adolescents, and together they account for more than 45 percent of all transferred tokens uttered by adults. Without them, the adult corpus only contains a relative amount of 20.8 transferred tokens per 1000 tokens, which increases the gap between adults and adolescents. This constitutes further support for the assumption that age plays an important role when it comes to the frequency of transferred lexical items in NG speech. A more fine-grained analysis of age differences among adults would be desirable for spoken NG. Unfortunately, the limited number of adult speakers in this part of the *DNam* corpus renders such an investigation problematic at the moment.

---

[24] A one-tailed test was performed because the alternative hypothesis was directed ("Adolescent speakers use more transferred tokens than adult speakers"). This is also why the tests for gender differences reported below are one-tailed as well (Alternative hypothesis: "Male speakers use more transferred tokens than female speakers").
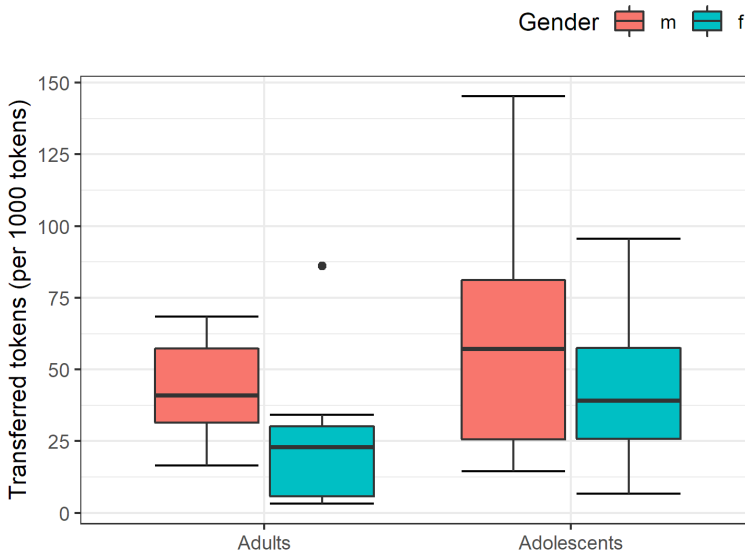
Figure 1: Individual relative frequencies of transferred tokens of male and female adults and adolescents

Now I turn to the frequency of use of transferred lexical items with respect to speaker gender. The corpus linguistic approach to the data shows that the male subcorpora contain relatively more transferred tokens than the female subcorpora in both age groups. The difference is 17 tokens per 1000 in the case of the adolescents and 22 tokens per 1000 in the case of the adults (see the values in square brackets in Tables 4 and 5, respectively). Since the adult subcorpus contains fewer transferred tokens overall, this means that the men's corpus contains almost twice as many transferred tokens as the women's corpus.[25] In both age groups the difference between the male and the female subcorpus is statistically significant according to the χ^2 test (adolescents: $\chi^2 = 131.587; p < 0.001; \varphi = 0.038$; adults: $\chi^2 = 53.961; p < 0.001; \varphi = 0.061$). The effect size is higher in the case of the adults.

So far the data seems to corroborate the assumption that male speakers use more transferred lexical items than female speakers. However, more than in the case of the age groups, the results of the analysis of individual frequencies for the gender groups deviate from the result of the corpus linguistic approach. Figure 1

---

[25]Importantly, excluding the three young speakers from Windhoek (two male, one female) from the sample of adults does not change the relative 2:1 ratio of other-language items between men and women. The remaining men use 110 transferred and 3169 other tokens, the remaining women use 136 transferred and 8395 other tokens ($\chi^2 = 35.993; p < 0.001; \varphi = 0.055$).

Table 4: Distribution of tokens over boys and girls		
	Tokens	
	Transferred [rel.]	Other
Boys	2190 [60.6]	33950
Girls	2392 [43.6]	52466

Table 5: Distribution of tokens over men and women		
	Tokens	
	Transferred [rel.]	Other
Men	233 [45.2]	4920
Women	221 [23.2]	9298

displays the individual frequencies of transferred tokens among male and female adults (boxplots on the left) and adolescents (boxplots on the right). It visualizes the dispersion of values between individuals within and among groups. Among the women dispersion is larger than among the men. On the one hand, three women used less than seven transferred lexical items per 1000 tokens, on the other hand the adult speaker who used the most (86.0) is also a woman, namely the young women from Windhoek mentioned above. Dispersion is even higher among adolescents. In this age group, the frequencies of female speakers are slightly more uniform than those of males. The first quartile (the lower border of the boxes in Figure 1) lies at just below 26 transferred tokens per 1000 tokens for both boys and girls but the third quartile of the boys lies at 81.1, exceeding the girls' third quartile by a margin of more than 23. Importantly, these individual results show that substantial differences exist also within the gender groups of each age group. Not all male and not all female speakers behave similar to one another. Furthermore, even though the median is higher by approximately 18 for male speakers in either age group, applying a one-tailed Wilcoxon rank-sum test provides results that are slightly above the significance level of 0.05 (adolescents: $W = 377; p = 0.079$; adults: $W = 37; p = 0.054$). That means, on an individual level, one cannot speak of a significant difference between male and female speakers in the same age group.

Concerning the relationship between speakers of the same gender in different age groups, it can be said that girls and women in the corpus behave less similar to each other than men and boys. In Figure 1, this is illustrated by the very small overlap between the boxes for the interquartile range of girls and women and it is reflected in a significant result of a one-tailed rank-sum test ($W = 186; p = 0.025$). The two groups behaving most similar are teenage girls and adult men who have almost identical medians (~40) and a similar interquartile range.

The analysis has already demonstrated that the quantity of transferred tokens is not a simple function of a single sociolinguistic variable. For the adolescents, I explored this further by considering an additional factor, namely the type of school attended by the speakers. I chose this variable (a) because school is the everyday social environment for students and this environment may influence their language use, and (b) because the schools attended by the speakers differ with respect to the role of German in the institution. That is, speakers are subject to different amounts of instruction in German and in other languages. The schools that speakers attended fall into three categories:[26]

1. The school receives funding from the Federal Republic of Germany, several subjects are taught in German: German Foreign School (*Deutsche Auslandsschule*, GFS).

2. The school offers the subject *Deutsch als Muttersprache* ('German as a mother language', DaM), no other subjects are taught in German: DaM school

3. The school does not offer any instruction in German: no-German-school. (Note, all students in the sample who attended this kind of school did, however, take private German lessons.)

The *Deutsche Höhere Privatschule* in Windhoek is the only school in Namibia belonging to the first category. Community members widely consider it an important institution for ("good") German in Namibia. Most students in the sample went to a school in the second category with some instruction in German, while a few students only took private German lessons (see Table 6). As mentioned earlier, the distribution of participants over gender and school type is not equal.

Table 6: Adolescent speakers per category of school

	Female	Male	Total
GFS	11	6	17
DaM	19	9	28
No German	2	4	6

---

[26]Previous tests for each individual school suggested that students of schools falling into one of these three categories behave more similarly to each other than to students from another group.
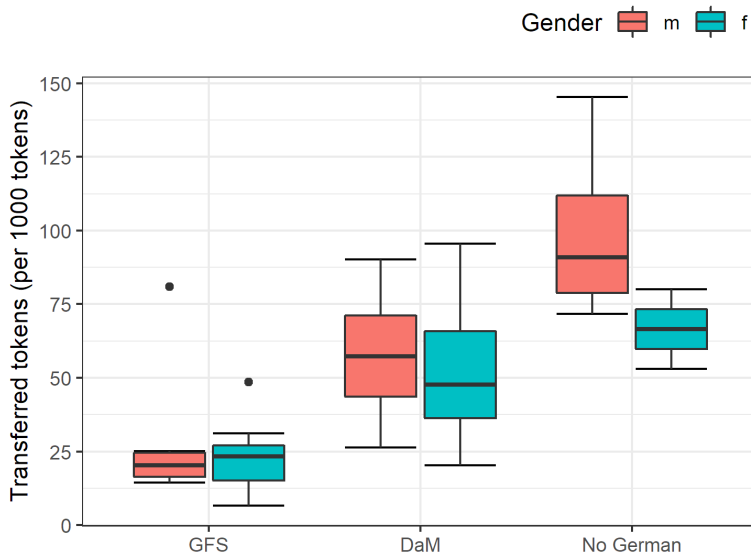
Figure 2: Individual relative frequencies of transferred tokens of male and female adolescents by school category

The results of analyzing the individual frequencies show an interesting pattern (Figure 2). There are significant differences between the three groups corresponding quite obviously to the status of German in the school.[27] At the same time, gender differences within groups are not significant. The median differs by less than three tokens per 1000 tokens (girls > boys) for the GFS and by less than ten tokens (boys > girls) for the DaM schools. For students attending schools without German instruction the gender differences are larger, with the boys' median being higher by a margin of 24. However, it has to be noted that this sample only contains four boys and two girls, which is also why the Wilcoxon rank-sum test could not provide a significant result. By contrast, the corpus linguistic approach does provide a significant difference for this sample ($\chi^2 = 25.054; p < 0.001; \varphi = 0.042$). For the other two school categories the corpus linguistic approaches results are as follows: There is no significant difference between boys and girls from the largest group, the DaM schools ($p = 0.997$). For the GFS school this approach indicates a significant difference ($\chi^2 = 25.431; p < 0.001; \varphi = 0.028$), but as Figure 2 shows this is due to a single male student who exceeds all of his fellow male students (and most of his fellow female students) by at least 50 transferred tokens per 1000. That is, the data really only suggests a difference between boys and girls in the smallest category, the no-German schools. Note that

---

[27] A one-tailed Wilcoxon rank-sum test gives the following results. GFS vs. DaM: $W = 428, p < 0.001$; DaM vs. No German: $W = 141, p = 0.004$.

this school category is also the only one in which there is an overproportional amount of male speakers. That is, the results for all boys and all girls presented above seem to be somewhat skewed by the fact that more than one in five boys in the sample goes to a no-German-school but only one in 16 girls. Therefore, exploring the additional sociolinguistic variable school type revealed that the findings concerning gender differences for adolescents should be treated with caution.

Unfortunately, a similar analysis for the adult speakers, which would focus on occupation and workplace, was not possible due to the small sample size of only 14 speakers. An observation suggesting that these aspects might well be relevant is that three women in the sample are school or pre-school teachers and two of them are among the women mentioned above who used very few transferred lexical items.

5.2 Proportion of donor languages

In addition to the frequency analysis of transferred tokens, I investigated how many of the transferred types and tokens are taken from English, Afrikaans and other Namibian languages. The results show that English and Afrikaans are by far the two most important donor languages in the data. 98.6 percent of the 1147 transferred types used by adolescents originate from one of these two languages. In the adult corpus, this applies to all except one of the 210 types. Thus, Bantu and Khoisan languages only play a minor role in the data as compared to the two Germanic languages, which is why I concentrate on the latter in the following. The columns *E* and *A* in Table 7 display the absolute frequencies of English and Afrikaans types and tokens in each of the four speaker groups, the percentages represent the proportion of English vs. Afrikaans types/tokens. These results show that, consistently, English is the dominant donor language in the data. In the corpora of all four speaker groups, the proportion of English types and tokens is higher than that of Afrikaans types and tokens. However, the ratios of English and Afrikaans differ somewhat between groups. English is most dominant in the girls subcorpus, followed by boys, then women and finally men. That is, with respect to all transferred types/tokens, adults used more material of Afrikaans origin than adolescents, and in both age groups male speakers used more material of Afrikaans origin than female speakers, with a larger difference between men and women.[28] For tokens (but not for types), all of these differences are significant.

---

[28]It is a different question how many Afrikaans-origin transferred tokens are used with respect to all tokens (not only with respect to all transferred tokens as above). This has to do with the

Table 7: Absolute frequency and proportion of English and Afrikaans types and tokens

	Types				Tokens			
	E	A	% E	% A	E	A	% E	% A
Boys	531	87	85.9	14.1	1782	363	83.1	16.9
Girls	614	81	88.4	11.7	2097	278	88.3	11.7
Men	98	37	72.6	27.4	156	76	67.2	32.8
Women	85	19	81.7	18.3	178	39	82.0	18.0

The dominance of English partly goes back to the use of dictionary-attested forms. For example, the 452 occurrences of the type *okay* account for almost 9 percent of all transferred tokens. However, as Table 8 displays, non-attested forms make up the majority of English-origin material in all groups. Even if only non-attested English-origin tokens/types are considered, these are still more frequent than Afrikaans-origin types and tokens. Table 8 also shows that, while the ratio of attested to non-attested English-origin types is rather similar across groups, the ratio of attested to non-attested *tokens* differs significantly in both age groups. Girls use about ten percent more attested tokens than boys ($\chi^2 = 47.496; p < 0.001; \varphi = 0.111$) and women about 14 percent more than men ($\chi^2 = 7.317; p = 0.007; \varphi = 0.148$).

Table 8: Distribution and proportion of English types and tokens in terms of dictionary attestation (att) or non-attestation (n-att)

	Types				Tokens			
	att	n-att	% att	% n-att	att	n-att	% att	% n-att
Boys	139	392	26.2	73.8	402	1380	22.6	77.4
Girls	153	461	24.9	75.1	682	1415	32.5	67.5
Men	22	76	22.5	77.6	46	110	29.5	70.5
Women	24	61	28.2	71.8	78	100	43.8	56.2

---

fact that adults use fewer transferred tokens overall. That is why women have a lower relative frequency (4.1) of Afrikaans tokens (per 1000 tokens) than girls (5.1) and boys (10.0). Despite the fact that men also use fewer transferred tokens than adolescents, they still have the highest proportion of Afrikaans-origin transferred tokens per 1000 tokens (14.7).

The overall tendency concerning the differences between male and female speakers with respect to the proportion of English and Afrikaans words is somehow reflected in the use of the most frequent Afrikaans-origin type, *net.* The word meaning 'just' or 'only', as in (8), occurs 199 times as a transferred token in the entire corpus.

(8) wir   wolltn   **net** essn   gehn   normal   un   das war
     we.NOM want.3PL.PRET just eat.INF go.INF normally and that be.3SG.PRET
     so fancy
     so fancy
     'We just wanted to go out for an ordinary dinner and that was so fancy.'
     [NAM092W1]

The relative frequency of *net* among male speakers is about twice as high as among female speakers in either age group. Importantly, *net* is also more frequent among male speakers relative to the Standard German alternatives *nur* and *bloß.* This can be shown by dividing the token counts for *net* by the sum of the token counts of all three words with similar meaning. The resulting proportion of *net* is 45.3 percent for boys and 30.4 percent for girls, 33.3 percent for men and 16.4 percent for women. The type is also more widespread among male speakers: Four out of six men and 17 out of 19 boys used it (89.5%), while only half of the eight women and 19 out of 32 girls (59.4%) used *net* at least once.

Much more could be said about the use of individual lexical items by different speakers. However, since this chapter has limited space, I now come to its final section for a summary and discussion.

6 Summary and discussion

Concerning the key aspect of the study, the quantity of transferred lexical items (other-language tokens, excluding instances of multi-word CS) in free conversations, the most obvious finding is that the phenomenon is ubiquitous in informal Namibian German. To some extent all speakers used transferred words, although the quantity varied substantially between individuals. Due to the perception of mixed and unmixed language in the NG community, a high frequency of transferred lexical items can be interpreted as a stronger deviation from the standard than a low frequency. The assumption that young speakers use more transferred lexical items than older speakers was generally supported by the data analysis. The results for the adults indicated that a more fine-grained analysis of language use by age would be desirable if the size of the data set allowed it. The results

concerning the assumption that male speakers used more transferred words than female speakers were less clear. The analysis suggests that the assumption is not true independent of other sociolinguistic variables. This became apparent by looking at gender and age, as well as at gender and school. For gender and age it was observed that, overall, teenage girls behaved rather similar to the group of adult men, while boys used the most and women the least relative amount of transferred tokens. Thus, within both age groups, male speakers used more transferred lexical items. However, whether this constitutes a significant difference depends on the statistical approach to the data. As opposed to age differences, gender differences were only significant with the broader corpus linguistic approach. Substantial intra-group differences were the reason for this. Furthermore, the analysis for gender and school indicated that gender differences in the group of adolescents may be an artifact of the sample composition. I observed that there are no significant gender differences for speakers of the DaM schools and that differences at the GFS go back to a single student. The differences are larger for speakers attending a school without instruction in German, with boys using more transferred lexical items than girls, but this subsample is also the smallest. Importantly, it was observed that, independent of gender, students who go to different types of schools (concerning the role of German) behave quite differently from one another. Students from the prestigious GFS, who receive the highest amount of German-language instruction, stood out as generally using very few transferred lexical items, while students with no subjects in German used the most. As the latter is the only group with more boys than girls the overall results for boys and girls are somewhat skewed in the direction that boys use more transferred tokens. Generally, this underlines the importance of a sophisticated quantitative analysis that takes into account different metadata and individual speaker behavior. The findings for the schools are also important in and of themselves. They strongly suggest that the frequency with which a speaker uses languages other than German in their everyday life influences their use of transferred lexical items. Moreover, since the schools have different reputations and orientations, the school a student attends likely says something about how important (Standard) German is for their parents. That is, these students are presumably confronted with different language ideologies in their home too, which they might have adopted.

In addition to the general frequency of transferred tokens, the role of donor languages was investigated. In the corpus data, English turned out to dominate quantitatively as a donor language, followed by Afrikaans, while other Namibian languages played only a minor role. The finding that English is more influential than Afrikaans is particularly noteworthy since it is at variance with previous

accounts of the role of donor languages in Namibian German, including the recent quantitative analysis by Zimmer (forthcoming). A reason for this may lie in the different data Zimmer and I used. Zimmer's study is based on Namibian German translations of "Wenker sentences", that is, on productions that are less spontaneous than informal conversations and might be subject to stylization.[29] Concerning the dominance of Afrikaans loanwords across all age groups in his data, Zimmer (forthcoming) himself suggests that this could be "an artefact of the design: maybe Afrikaans words are considered particularly salient and are used in the translations to emphasise the deviance of Namdeutsch from European German". This assumption is corroborated by several interview statements from community members in the *DNam* corpus. When asked about vocabulary they perceive as typical for NG, speakers predominantly mentioned terms originating from Afrikaans, such as *braaien* (see §2), *net* (see §5.2), *mooi* (see example 3), *pad* ('road'), *lekker* (in the sense of 'good'/'pleasant'). Yet, spontaneous language in the *DNam* corpus turned out to be influenced by English to a much larger extent than by Afrikaans. An explanation for this apparent mismatch might be that words from Afrikaans are more salient because they are older, since Afrikaans was more important than English until Namibian independence in 1990. Speakers may still know many Afrikaans words because these used to be more common in the past, therefore perceive them as salient, and accordingly put them to use in translations of "Wenker sentences". At the same time, Afrikaans words might be declining in actual NG language use, which would explain the results of this study. This might also explain the fact that younger speakers used even fewer transferred words from Afrikaans than older speakers in their informal conversations. It would not, however, account for the differences between male and female speakers in both age groups. Still, the salience of Afrikaans might play a role here as well. Afrikaans words are perceived as indexical of a traditional Namibian German identity, which, as I have argued above, seems to have connotations of stereotypical masculinity. Therefore, sounding "typically Namibian" might overlap with sounding "typically male" and speakers who seek to construct a traditional male identity might do so by using features also indexical of a traditional Namibian German identity. By contrast, English is presumably rather associated with modern Namibia and a globalized, English-speaking world. Thus, for speakers orienting themselves away from traditional views and structures in the community, English might be more attractive. This could be a reason why

---

[29]Cf. also Radke's (2017: 116) finding that the most frequent transferred items in highly stylized newspaper commentaries in the German-language Namibian newspaper *Allgemeine Zeitung* are predominantly of Afrikaans origin.

young female community members are particularly inclined towards English-origin lexical material.[30] It would be necessary to further investigate these hypotheses and the assumptions I made. One could, for example, study the community attitudes towards English and Afrikaans on the one hand and towards (traditional) gender roles on the other and explore whether these attitudes are reflected in speakers' language use. This study has indicated that there might be an interesting sociolinguistic dynamic at work.

Although I have focused on the aspects of gender and age, this study has also indicated that these are not the only aspects that play a role for the use of transferred lexical items in NG. The findings concerning schools suggest that everyday language use is of importance. This also raises the question of differences between urban and rural dwellers because Namibian German farmers often speak Afrikaans with black farm laborers and with farmers belonging to the Afrikaans-speaking community, whereas in cities English prevails as the language of business. Aside from the place of work and living, language use with family and friends is certainly another aspect to focus on in future studies. The role and interaction of various sociolinguistic variables such as those discussed here in the use of transferred lexical items could further be investigated using a multifactorial statistical model.

I will conclude this chapter with some remarks concerning its methodological aspects. I think, with this study, I have made a case for a careful approach to corpus data. When studying a phenomenon in a corpus that contains productions by a number of speakers, researchers should not only look at aggregated totals but also at the individual behavior of all speakers. Only then it is possible to assess how widespread the phenomenon in question is and how strongly its occurrence varies within the corpus (cf. Gries 2010: 274). Lastly, as a prerequisite for the analysis, I have presented my approach to categorizing other-language material in the corpus with the help of a simple annotation system. Note again that the results presented in this chapter concern a subset of the other-language material in the corpus selected on the basis of that annotation, excluding all tokens treated as multi-word code-switches. Certainly, studies based on different selections of other-language data are conceivable. As the annotations will be available for users of the *DNam* corpus in the future, it will be possible for researchers interested in language mixing phenomena in NG to choose their own set of other-language data. I hope that this chapter has helped to stimulate interest in such further research and look forward to its results.

---

[30]It was also observed that female speakers use significantly more dictionary-attested English-origin tokens than male speakers in the same age group. That is, in this respect the speech of female speakers appears to be closer to the prestige variety Standard German.

Acknowledgements

References

Auer, Peter. 2011. Dialect vs. standard: A typology of scenarios in Europe. In Bernd Kortmann & Johan van der Auwera (eds.), *The languages and linguistics of Europe: A comprehensive guide* (The World of Linguistics 1), 485–500. Berlin: de Gruyter. DOI: 10.1515/9783110220261.485.

Baker, Paul. 2014. *Using corpora to analyze gender.* London: Bloomsbury. DOI: 10.5040/9781472593559.

Böhm, Michael-Anton. 2003. *Deutsch in Afrika: Die Stellung der deutschen Sprache in Afrika vor dem Hintergrund der bildungs- und sprachpolitischen Gegebenheiten sowie der deutschen auswärtigen Kulturpolitik* (Duisburger Arbeiten zur Sprach- und Kulturwissenschaft 52). Frankfurt: Lang.

Calude, Andreea Simona, Steven Miller & Mark Pagel. 2017. Modelling loanword success: A sociolinguistic quantitative study of Māori loanwords in New Zealand English. *Corpus Linguistics and Linguistic Theory* 16(3). 29–66. DOI: 10.1515/cllt-2017-0010.

Coates, Jennifer & Pia Pichler (eds.). 2011. *Language and gender: A reader.* 2nd edn. Malden: Wiley-Blackwell.

Deumert, Ana. 2009. Namibian Kiche Duits: The making (and decline) of a Neo-African language. *Journal of Germanic Linguistics* 21(4). 349–417. DOI: 10.1017/S1470542709990122.

Eckert, Penelope. 2003. Language and gender in adolescence. In Janet Holmes & Miriam Meyerhoff (eds.), *The handbook of language and gender*, 381–400. Malden: Blackwell. DOI: 10.1002/9780470756942.ch16.

Eckert, Penelope. 2011. Gender and sociolinguistic variation. In Jennifer Coates & Pia Pichler (eds.), *Language and gender: A reader*, 2nd edn, 57–66. Malden: Wiley-Blackwell.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41(1). 87–100. DOI: 10.1146/annurev-anthro-092611-145828.

Eckert, Penelope & Sally McConnell-Ginet. 1999. New generalizations and explanations in language and gender research. *Language in Society* 28(2). 185–201. DOI: 10.1017/S0047404599002031.

Eisikovitz, Edina. 2011. Girl-talk/boy-talk: Sex differences in adolescent speech. In Jennifer Coates & Pia Pichler (eds.), *Language and gender: A reader*, 2nd edn, 38–48. Malden: Wiley-Blackwell.

Gregersen, Johanna & Nils Langer. 2021. Assessing language contact: Linguistic purism and North Frisian. In Christian Zimmer (ed.), *German(ic) in language contact: Grammatical and sociolinguistic dynamics*, 159–186. Berlin: Language Science Press. DOI: 10.5281/zenodo.4954485.

Gries, Stefan Th. 2010. Useful statistics for corpus linguistics. In Aquilino Sánchez Pérez & Moisés Almela Sánchez (eds.), *A mosaic of corpus linguistics: Selected approaches* (Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation 66), 269–291. Frankfurt: Lang.

Haspelmath, Martin. 2009. Lexical borrowing: Concepts and issues. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 35–54. Berlin: de Gruyter. DOI: 10.1515/9783110218442.35.

Kellermeier-Rehbein, Birte. 2015. Namslang – Deutsche Jugendsprache in Namibia? In Kerstin Runschke & Corinna Peschel (eds.), *Sprachvariation und Sprachreflexion in interkulturellen Kontexten*, 41–62. Frankfurt: Lang.

Kellermeier-Rehbein, Birte. 2016. Varietäten der deutschen Sprache in Namibia. In Thomas Stolz, Ingo H. Warnke & Daniel Schmidt-Brücken (eds.), *Sprache und Kolonialismus: Eine interdisziplinäre Einführung zu Sprache und Kommunikation in kolonialen Kontexten*, 213–234. Berlin: de Gruyter. DOI: 10.1515/9783110370904-010.

Labov, William. 2006. *The social stratification of English in New York City*. 2nd edn. Cambridge: Cambridge University Press.

Macaulay, Ronald K. 1977. *Language, social class, and education*. Edinburgh: Edinburgh University Press.

Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, Carol. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.

Nöckler, Herbert C. 1963. *Sprachmischung in Südwestafrika*. München: Hueber.

Poplack, Shana. 2004. Code-switching. In Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier & Peter Trudgill (eds.), *Sociolinguistics: An international handbook of the science of language and society*, 2nd edn (Handbooks of Linguistics and Communication Science 3.1), 589–596. Berlin: de Gruyter. DOI: 10.1515/9783110141894.1.4.589.

Poplack, Shana. 2018. *Borrowing: Loanwords in the speech community and in the grammar.* New York: Oxford University Press. DOI: 10.1093/oso/9780190256388.001.0001.

Poplack, Shana, David Sankoff & Christopher Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26(1). 47–104. DOI: 10.1515/ling.1988.26.1.47.

Pütz, Joe. 1982. *Dickschenärie.* Windhoek: Camel Graphics.

Radke, Henning. 2017. "Lekker dieses namtal zu lesen. Ich hou davon.": Namdeutscher Sprachgebrauch in namibischer Onlinekommunikation. *Germanistische Mitteilungen* 43(2). 109–132. DOI: 10.33675/GM/2017/2/2.

Radke, Henning. 2021. Language contact and mixed-mode communication: On ingroup construction through multilingualism among the German-Namibian diaspora. In Christian Zimmer (ed.), *German(ic) in language contact: Grammatical and sociolinguistic dynamics*, 127–158. Berlin: Language Science Press. DOI: 10.5281/zenodo.4954483.

Schmidt, Thomas, Wilfried Schütte & Jenny Winterscheid. 2015. *cGAT: Konventionen für das computergestützte Transkribieren in Anlehnung an das Gesprächsanalytische Transkriptionssystem 2 (GAT2).* http://agd.ids-mannheim.de/download/cgat_handbuch_version_1_0.pdf (17 May, 2021).

Schmidt-Lauber, Brigitta. 1998. *Die verkehrte Hautfarbe: Ethnizität deutscher Namibier als Alltagspraxis* (Lebensformen 10). Berlin: Reimer.

Shah, Sheena. 2007. German in a contact situation: The case of Namibian German. *eDUSA* 2(2). 20–45. http://eprints.soas.ac.uk/23364/1/2007_namibian_german.pdf (17 May, 2021).

Shah, Sheena & Marianne Zappen-Thomson. 2018. German in Namibia. In Corinne A. Seals & Sheena Shah (eds.), *Heritage language policies around the world* (Routledge Studies in Sociolinguistics 15), 128–147. London: Routledge. DOI: 10.4324/9781315639444-9.

Tagliamonte, Sali A. 2012. *Variationist sociolinguistics: Change, observation, interpretation* (Language in Society 40). Malden: Wiley-Blackwell.

Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1(2). 179–195.

Wiese, Heike. 2020. Language situations: A method for capturing variation within speakers' repertoires. In Yoshiyuki Asahi (ed.), *Methods in dialectology XVI* (Bamberg Studies in English Linguistics 59), 105–117. Frankfurt: Lang.

Wiese, Heike & Yannic Bracke. 2021. Registerdifferenzierung im Namdeutschen: Informeller und formeller Sprachgebrauch in einer vitalen Sprechergemeinschaft. In Csaba Földes (ed.), *Kontaktvarietäten des Deutschen im Ausland* (Beiträge zur Interkulturellen Germanistik 14), 273–293. Tübingen: Narr.

Wiese, Heike, Horst J. Simon, Marianne Zappen-Thomson & Kathleen Schumann. 2014. Deutsch im mehrsprachigen Kontext: Beobachtungen zu lexikalisch-grammatischen Entwicklungen im Namdeutschen und im Kiezdeutschen. *Zeitschrift für Dialektologie und Linguistik* 81(3). 274–307.

Wiese, Heike, Horst J. Simon, Christian Zimmer & Kathleen Schumann. 2017. German in Namibia: A vital speech community and its multilingual dynamics. In Péter Maitz & Craig A. Volker (eds.), *Language contact in the German colonies: Papua New Guinea and beyond* (Language & Linguistics in Melanesia, special issue), 221–245. https://www.langlxmelanesia.com/LLM%20S%202017_Wiese%20et%20al.pdf (17 May, 2021).

Wolfram, Walt. 1969. *A sociolinguistic description of Detroit negro speech*. Washington: Center for Applied Linguistics.

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. DOI: 10.1177/1367006914521699.

Zimmer, Christian. 2019. Deutsch als Minderheitensprache in Afrika. In Joachim Herrgen & Jürgen Erich Schmidt (eds.), *Sprache und Raum – Deutsch: Ein internationales Handbuch der Sprachvariation* (Handbooks of Linguistics and Communication Science 30.4), 1176–1190. Berlin: de Gruyter. DOI: 10.1515/9783110261295-046.

Zimmer, Christian. Forthcoming. Linguistic variation and age of speakers in Namibian German: Loan word usage in "Wenker sentences". In Hans C. Boas (ed.), *German abroad: Comparative perspectives on language contact*. Leiden: Brill.

Zimmer, Christian, Heike Wiese, Horst J. Simon, Marianne Zappen-Thomson, Yannic Bracke, Britta Stuhl & Thomas Schmidt. 2020. Das Korpus *Deutsch in Namibia* (*DNam*): Eine Ressource für die Kontakt-, Variations- und Soziolinguistik. *Deutsche Sprache* 48(3). 210–232. DOI: 10.37307/j.1868-775X.2020.03.03.