

**Note: This is not the final published version of the article, but the ‘accepted’ version. To quote the article, please refer to the published version.**

Peterson, Elizabeth, Ylva Biri & Johanna Vaattovaara 2022: Grammatical and social structures of English-sourced swear words in Finnish discourse. –Martín-Solano, R. & San Segundo, R. (Eds.), *Corpus linguistics and Anglicisms: Corpus-Aided Research into the Influence of English on European Languages*, 49–70. Peter Lang Publishing.

## **Grammatical and social structures of English-sourced swear words in Finnish discourse**

Elizabeth Peterson (University of Helsinki), Ylva Biri (University of Helsinki), Johanna Vaattovaara (Tampere University)

In this chapter, we investigate the use of English-sourced swear words compared to heritage Finnish swear words in Finnish discourse. The investigation was conducted in two phases: 1) a corpus investigation 2) a perception study using both oral and written sample materials in an online survey (i.e., reaction tasks). With our investigation of the corpus data, we note the relationship of English-sourced swear words compared to heritage Finnish swear words in an online forum, focusing on distribution within the clause, any discernable discourse patterns, and the grammatical function of the swear words. Phase 2 involves social and regional categories. Based on survey responses from 446 Finnish speaking respondents, the results indicate that each English-sourced swear word investigated is integrated differently into Finnish discourse and society. The English borrowings *shit* and *fuck* are notable examples of this integration. Some of the borrowings show clear associations with the Helsinki urban region, some with age, and others with gender.

swearing; language contact; English; Finnish; perception; language attitudes; CMC

### **1. Introduction**

Previous research on English-sourced pragmatic borrowings in Finnish indicates that these borrowings, for example *pliiis* ‘please’ and *oh my god*, tend to index social meanings such as urbanicity, youth and also women (Peterson 2017; see also Peterson and Vaattovaara 2014). When it comes to the use of English-sourced swear words in Finnish discourse, this leads to interesting questions. In many societies, swear words are associated with men (e.g., McEnery 2006 on English, but see Stapleton 2003, 2010 for counterevidence). In Finland, pragmatic borrowings from English are stereotypically connected with women. What does this mean, then, in terms of the borrowing of English swear words into Finnish? This two-pronged investigation of swear words in Finnish compares the borrowings *damn*, *fuck* and *shit* to the Finnish equivalents *perkele/hitto*, *vittu* and *paska*, respectively. The selection of lexemes was based on frequency as well as semantic/functional equivalents of the Finnish and the English-sourced words.

To gain insights into both grammatical and social structures of swear word borrowings from English into Finnish, we first made a general investigation into the use of English swear words compared to Finnish swear words in Finnish computer-mediated discourse, using Finland’s biggest online forum, Suomi24, available as a corpus from Fin-Clarín. With our investigation of the corpus data, we were interested in the use of English-sourced swear words compared to heritage Finnish

swear words. We focused on their distribution within a clause, any discernable discourse patterns, and the grammatical function of the forms, such as word class.

The second stage of our investigation viewed the perceptual climate of swear-word usage, investigating swear words in connection to social factors. Making use of a modification of a classic audio matched guise test (Lambert 1967) and a self-report task, we received responses to an online survey from 446 speakers of Finnish from across Finland. The overall results show that borrowed swear words tend to be associated with younger people, the Helsinki urban area, and women, although there are nuances among specific borrowings.<sup>1</sup>

## 2. Background

As with English swear words, Finnish swear words relate to the semantic domains of taboo subjects, including excrement, sexual slurs or words with religious meaning. Thus, the word *paska* means ‘shit’, whereas both *hitto* and *perkele* come from a religious background. *Hitto* stems from *hiisi*, referring to sacred places, including burial sites, and the mythical creatures who live there (Kulonen 1990). *Hitto* in Finnish seems to correspond to the pragmatic force of *damn* in English (Vaattovaara & Peterson, under review). Meanwhile, *perkele* means ‘devil,’ although not exclusively in the Christian sense, as the root of *perkele* is the Baltic word for ‘thunder’ and ‘thunder god,’ retained in modern Lithuanian as *perkūnas* (Kulonen 1990). Semantically, *vittu* shares a referent with the English word *cunt*, although as a swear word the usage differs in morphosyntactic and pragmatic function. Namely, whereas *cunt* mainly refers to people, *vittu* can be inflected to refer to various objects and exists also in the verb form *vituttaa* ‘to be angry.’ *Vittu* thus has a wider productivity, comparable to that of *fuck* in English (Hjort 2017).

Collectively, English-sourced swear words in Finnish are prime examples of what has been referred to as *pragmatic borrowing* (Andersen 2014), or in other words luxury loans (Bloomfield 1933) and discourse styles that signal speaker attitudes and stylistic choices rather than adding propositional meaning to an utterance. Importantly, pragmatic borrowings tend to be in variation with heritage forms in the receiving language. Finnish is by no means impoverished when it comes to swear words and expletives, which raises an important question: why incorporate swear word borrowings from English? What are the English swear-word borrowings accomplishing in Finnish discourse that Finnish heritage forms do not? Or, put another way: what is the motivation for the borrowing of English swear words?

Previous work on pragmatic borrowings in Finnish has demonstrated that such borrowings, for example the adjective *jees* ‘okay, good’ (Pyykkö 2017) and the lexical politeness marker *pliis* (Peterson and Vaattovaara 2014; Peterson 2017), tend to be associated with urban settings, a global orientation, and youth language. While such borrowings often appear to be perceptually associated with women, the forms are, unsurprisingly, not excluded to women (see Peterson and Vaattovaara 2014). Findings from our previous work has informed the study reported on here. With the current study we aimed to observe, for example, if the English-sourced swear words we investigated appear to be more available for people in urban settings in Finland than in rural settings. We also aimed to investigate if the English-sourced swear words are perceived differently according to age, gender or education of the respondents. These aims were addressed through the modified oral matched guise test (see section 4).

With this study, we also broaden the scope of our investigation to explore grammatical integration of the borrowed forms. Our previous work on pragmatic borrowings has focused primarily on discourse particles, and, as such, we have not had concerted opportunity to investigate

---

<sup>1</sup> The detailed reports of the survey results are published elsewhere (Vaattovaara and Peterson 2019; Vaattovaara and Peterson, under review).

integration into Finnish morphology. The corpus portion of our study enabled us to gain empirical evidence about how Finnish people use English-sourced swear words in Finnish syntax and morphology.

### 3. Phase 1: Corpus study

With this portion of investigation, the main aims were to gain insight into the swearing patterns of Finnish discourse, especially with regard to pragmatic borrowings from English. Specifically, we analyzed the corpus data with the following questions in mind:

RQ1) How do the use of heritage swearing forms compare to borrowed forms in frequency?

RQ2) How do the grammatical patterns of borrowed swear words differ from those of heritage forms?

RQ3) What are the levels of integration of the borrowings morphologically and semantically?

With regard to RQ1, we wished to observe if there are grammatical and frequency distinctions in the use of the borrowed versus heritage forms in the CMC-based corpus data (described later in this section). By this, we mean that Finnish heritage forms might, for example, be used more frequently as subjects or objects, whereas English swear word borrowings might be used primarily in a discourse particle function. With RQ2, the observations were mostly qualitative: what kind of creative or performative uses do we find with regard to the use of English swear word borrowings? With RQ3, we wished to observe the level of integration, if any, English-sourced swear words have in the rich morphological system of Finnish. The structural elements of Finnish also have a bearing on RQ1 and RQ2.

#### 3.1 The corpus data

To gain naturally-occurring, empirical data on swear words in everyday Finnish discourse, we turned to the Suomi24 corpus, available as a corpus from Fin-Clarín.<sup>2</sup> The corpus is created from the Suomi24 online forum, one of Finland's largest online social networking sites. Suomi24.fi hosts a general discussion forum with topics divided into several subfora ranging from pastime interests to politics and society. The forum allows participation both anonymously and through login. The forum reports nearly even distribution between genders, with 52 percent female and 48 percent male [sic] users (Aller Media ltd 2019). While communication on internet forums such as Suomi24 is, technically, asynchronous written Finnish, it is a format that also is representative of everyday, colloquial Finnish.

The Suomi24 corpus is a text corpus that is part of the Fin-Clarín collection. Lagus et al. (2016:10) observe that besides the prevalence of trolling or deliberate aggravation, subfora or discussion threads related to small towns, unemployment and alcohol often take a turn for the more inappropriate, thus requiring moderator intervention. As the corpus does not contain comments removed by forum moderators, swear words with explicitly aggressive functions may not be represented by the corpus. However, the corpus still contains instances of swearing for various other pragmatic functions, such as emphasis or expressing positive or negative emotions. In other words,

---

<sup>2</sup> <https://www.kielipankki.fi/language-bank/>

unless it is deemed as aggressively directed toward certain segments of the user population, swearing is allowed and used freely on Suomi24.

For this study, we used the publicly available Suomi24 2016H2 corpus (Aller Media ltd. 2014) with 2.66 billion word-tokens of forum entries posted between 2001 and 2016. A random sample of 150 instances per each swear word (see Table 1) was collected in June 2017 using the Korp concordance search interface, developed by Språkbanken (Borin et al. 2012). Due to the high frequency of the word  *fucking* , it was considered separately from the base form  *fuck*  and studied in a separate query with a random sample of 150. Other inflections of  *fuck* , a total of 37 instances, were studied separately.

Word	Total hits	PPM
<i>paska</i> ‘shit’	729,413	273.89
<i>vittu</i> ‘fuck’	441,207	165.67
<i>saatana</i> ‘satan / hell’	363,577	136.52
<i>perkele</i> ‘damn / hell’	138,946	52.17
<i>shit</i>	14,668	5.51
<i>fuck</i> (base form)	5,846	2.20
<i> fucking </i>	3,525	1.32
<i>damn</i>	1,616	0.61
<i>fuck*</i> (other forms)	37	0.01

Table 1: Swear word lemma frequency in Suomi24 corpus

To assess the integration of the swear word tokens, the sample concordances of the swear words were analyzed in terms of the position of the swear word within a clause. The degree of syntactic integration was manually classified based on the word’s grammatical form (inflected for agreement, if relevant) and its position as part of or outside of a clause. Such distinctions were based on grammatical constructions, visual information such as punctuation as well as pragmatic factors such as utterance types (e.g. exclamations). The heritage Finnish and the English-sourced swear words were identified as sharing a total of four categories, shown in Table 2. English-sourced swear words were further given a fifth category, “code-switching,” for cases representing a full clause of English text. The data was categorized according to the parameters demonstrated in Table 2.

Category	Definition	Examples
isolated	A self-contained clause, e.g. an exclamation.	<i><u>Damn</u></i> <i><u>Saatana!!!!</u></i>
particle	Within a clause, but does not fill a grammatical function, or has not been inflected in agreement to fit a syntactic role in Finnish. Pragmatically, an interjection.	<i>jossain vaiheessa sitä vaa ajattelee et <u>damn</u>, jos en ois eronnu ni en ois tavannu tätä upeeta ihmistä</i> ‘at some point one just thinks that <u>damn</u> , if I hadn’t divorced, I wouldn’t have met this amazing person’
integrated	Fills syntactic function in clause, or is inflected according to grammar rules; e.g. modifier inflected to genitive.	<i>ihanku se ois joku <u>damn</u> ihme</i> ‘as if it’s some <u>damn</u> miracle’ <i>kun tuo <u>saatanan</u> ryssäniivana taas hyökkäsi maahamme</i> ‘when that satan-GEN Russian Ivan attacked our country again’
clustered	Occurs in a sequence with repetition of swear word(s), may be inflected	<i><u>vitun paska saatana</u></i> ‘fuck-GEN shit satan’

code-switching	English-sourced swear word is part of a longer chunk of English.	<i>Niin se maailma muuttuu heh, you can't do a <u>damn</u> thing about it</i> 'That's how the world changes heh, you can't do a damn thing about it.'
----------------	--	---

Table 2: Classification system for swear word tokens in Suomi24 data

Occurrences of swear words that did not have the semantic meaning or pragmatic function of swearing were marked as unapplicable; such cases included, for example, instances of *saatana* in its religious sense of 'satan' or metalinguistic commentary (e.g. *damn on englantia ja tarkoittaa suomenkielistä sanaa hitto* 'damn is English and means the Finnish word hitto'). Swear words occurring in famous quotes or established sayings were also marked unapplicable (e.g. *Frankly my dear, I don't give a damn* and *No shit, Sherlock*).

### 3.2 Results of corpus analysis

The results are presented in terms of research questions, beginning with the general question of frequency of forms.

*RQ1) How do the use of heritage swearing forms compare to borrowed forms in frequency?*

The corpus analysis showed that English-sourced swear words in the Suomi24 forum constituted a small overall portion of swearing in Finnish online discourse. As can be seen in the Table 1, the number of English-sourced swear words occurred at a ratio of 1:1044 compared to the Finnish swear words. This finding indicates that English-sourced swear words are not taking over the territory of heritage Finnish swear words. Rather, it appears that English-sourced swear words are used for particular social and pragmatic functions that are not perceived by users as being carried out by the heritage forms.

*RQ2) How do the grammatical patterns of borrowed swear words differ from those of heritage forms?*

Due to the robust morphological system of the Finnish language, we suspected that English-sourced swear words would be favored for use in the corpus as discourse particles. This is because discourse particles do not require morphological or clausal integration (see Müller 2005). That is, rather than individual speakers opting to integrate English swear word borrowings into Finnish morphology and syntax, we expected to find evidence of English-sourced forms being favored in discourse-particle functions. However, this was not entirely the case—at least not for all the borrowed swear words investigated. As illustrated in Figure 1, the words *shit*, *fuck* and *fucking*—that is, the most frequent of the English-sourced forms—are integrated into the sentence 48.7 percent, 10.7 percent and 34 percent of the time, respectively. Meanwhile, *damn* is used as a discourse particle 22 percent of the time and integrated only 2.7 percent of the time.

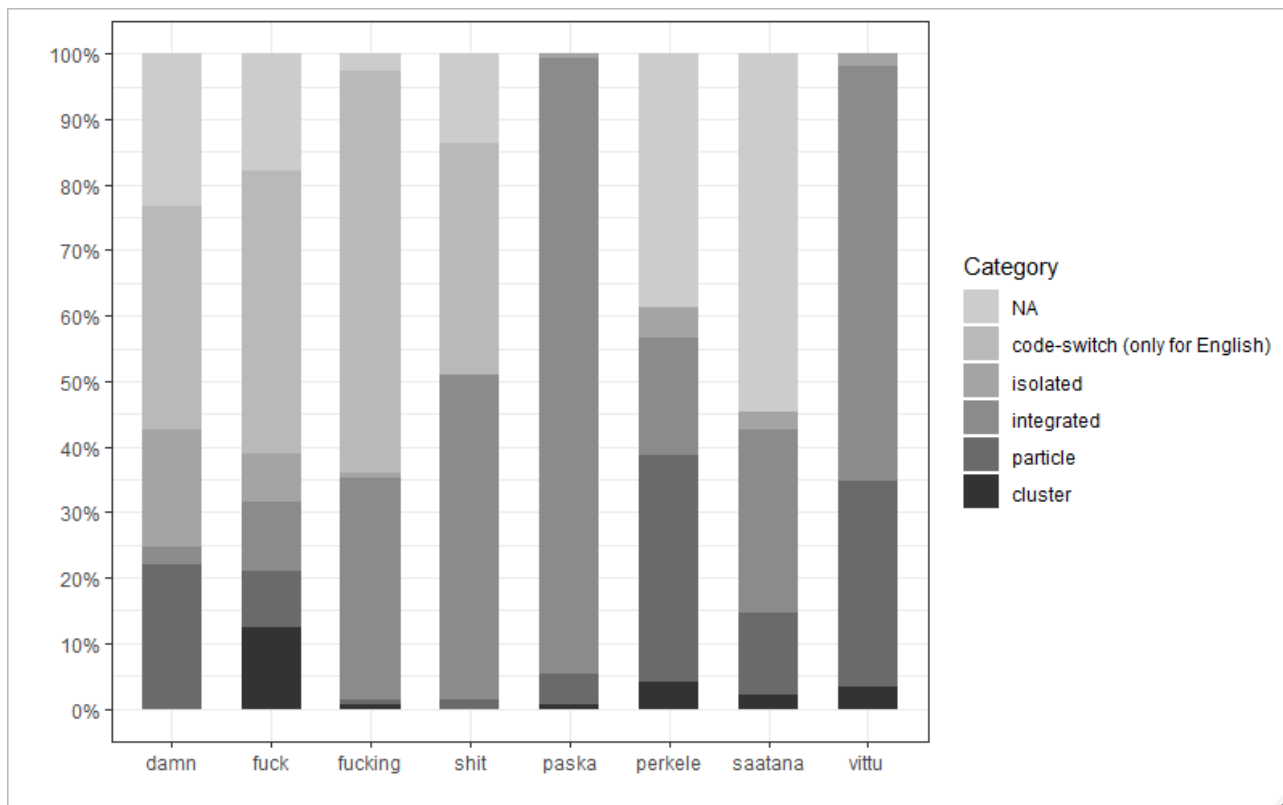


Figure 1. Distribution of patterning classes for English-sourced and Finnish-heritage swear words

Compared to the other borrowed words, the proportions of *damn* as a particle or as an isolated utterance is noticeably higher than for the other more popular words. This suggests a perception of *damn* seen as less suitable for inflection or more suitable for peripheral discourse particle use. Contrastingly, the frequently occurring words *fuck* and *shit* are often integrated into the sentence, resulting also in their inflection to different grammatical cases to meet the requirements of Finnish grammar (see Beers Fägersten 2014 and 2017 for similar findings in Swedish). *Shit* and *fuck* show a high level of integration (see also Vaattovaara & Peterson 2019), exemplified by the following:

- (1) *kyllä on todellista **shittiä***  
 yes be.3SG true-PTV shit-PTV  
 ‘Yes it is truly shit’
- (2) *... **shitteihin** asti.*  
 shit-PL.ILL upto  
 ‘all the way to the shits’
- (3) *sinä **fucking** narsistipaska!*  
 2SG fucking.NOM narcissist.shit.NOM  
 ‘You fucking narcissist shit!’

We observed that integration occurs also without inflection, as some instances of integration exemplify grammatical borrowing. In example (3), both the head and the modifying swear word are in the nominal case. However, example (4) illustrates an English borrowing functioning as modifier without the grammatically expected case inflection.

- (4) *takaisin siihen **fucking** tiskiän päin*

back 3SG.ILL fucking.NOM bar.ILL towards  
'Back towards the fucking bar'

In other words, the English-sourced modifier remains in base form just as it would in English, whereas a Finnish adjective would be in the genitive form.

To summarize, English-sourced swear words can be inflected to match Finnish morphology and be assigned the same grammatical functions as Finnish words. The integration of English-sourced swear words can be best illustrated through *fuck* as a noun, which was found to be used very flexibly in our data sample. In our data, it was used with six different nominal case endings: genitive (*fuckin*), partitive (*fuckkia*), plural partitive (*fuckeja*), inessive (*fuckissa*), elative (*fuckista*) and illative (*fuckiin*). As the less frequent word *damn* indicated little integration, our findings suggest that grammatical integration of English-sourced swear words occurs but is limited to the most frequent words.

*RQ3) What are the levels of integration of the borrowings morphologically and semantically*

When integrated into Finnish, English-sourced swear words can bring pragmatic or idiomatic meanings or grammatical combinations that do not exist in English. While representing only a small portion of integrated swear words, such instances of localized meanings illustrate integration both on grammatical and semantic levels. Example (5) demonstrates typical integration, where the case system of Finnish grammar requires the swear word and its modifier to be inflected. Here, however, the borrowing is not limited just to the word *shit* but extends to the entire phrase *deep shit*.

- (5) *alkaa taas olemaan sen verran diippiä shittiä, että ois* [standard Finnish=*olisi*] *paree*  
[standard Finnish=*parempi*] *lopettaa.*  
begin.INF again be.INF.III 3SG.INANIM.GEN amount-GEN deep-PTV shit-PTV that be.3SG.COND  
better end.INF  
'It starts to be such deep shit again, that it would be better to stop.'

The word *diip*, a localized spelling of 'deep,' is inflected to agree with *shit*, rendering the English-sourced idiom *deep shit*. This expression, coupled with the nonstandard, colloquial spelling choices in the utterance, point toward an overall pragmatic effect of informality and localness.

As noted above, the word *fuck* is often grammatically integrated; however, this is not limited to inflectional cases, as we also found instances of derivational productivity. Our query of the inflected word forms of *fuck*, excluding *fucking*, yielded several instances of inflected forms, representing integrated words. Example (6) is a noteworthy instance of *fuck* inflected into a nominalized verb and used as part of a Finnish vulgar word for 'nitpicker,' *pilkunnussija*. Literally, this term is translatable as 'comma fucker,' but the perceived meaning is someone who is overly pedantic or nitpicky. In (6), the word *pilkku* 'comma' is retained in Finnish, with the English borrowing replacing the *nussija* 'fucker':

- (6) *Miksi sinua kiinnostaa siis pilkkujen fuckkaaminen enää, jos kommunikaation oleellinen tarkoitus: ymmärrys*  
why 2SG.PTV interest-INF that.is comma-PL.GEN fuck-INF.IV anymore if communication-GEN essential.NOM purpose.NOM : understanding.NOM  
'Why are you interested in fucking commas [=being nitpicky] anymore, if the essential purpose of communication: understanding'

In other words, the Finnish equivalent is partially replaced by an English borrowing, resulting in a mixed-language compound while retaining the meaning of *pilkunnussija*.

In the data, lemmas derived from *fuck* also appear in other novel usages, possibly to fill a function or association that is not felt to be sufficiently expressed by heritage swear words. In the following, *fucking* is combined with Guggenheim into a slur against the proposed building of a Guggenheim Museum in Helsinki, which caused debate and resistance in the year 2014.

- (7) *voihan ne rikkaat itse maksaa sen fuckkenhaimin*  
can.3SG-CLITIC 3PL.INANIM self pay.INF.3SG 3SG.GEN fuckenhaim-GEN  
'the rich could pay for the Fuckenheim [= a blend of *fuck* and Guggenheim museum] themselves'

Similarly, in (8), *fuck* is combined with the derivational suffix *-isti* ('-ist'), creating a term that, in this case, refers to a member of the social welfare state.

- (8) *Olen itsekin suomalainen nuori, mutta en fuckisti*  
be.1SG self-CLITIC Finnish.NOM youth.NOM but not.1SG fuckisti  
'I am myself a Finnish youth, but I am not a fuckisti.'

In summary, English-sourced swear words are integrated into Finnish morphology through both inflection and derivation. Instances of localized meanings include both the borrowing of English meanings and idioms and combinations with Finnish semantics. While these instances are limited to the most frequent English borrowings—*shit* and *fuck*—the findings suggest that these words are most subject to grammatical and pragmatic integration.

#### 4. Phase 2: Perception study

Phase 1 of the study revealed a few key findings. For one, it became clear that, according to this data set, swear word borrowings from English are marginal in use compared to heritage Finnish swear word forms. At the same time, two English swear word borrowings in particular, namely *shit* and *fuck*, showed a surprisingly high level of integration into Finnish grammar and discourse. In addition, one swear-word borrowing patterned differently from the others: *damn* was more frequent as a particle than the other English-sourced swear words we investigated. Thus, while phase 1 of the study was a necessary stage in offering an overview of the general patterns of swear-word use in Finnish, it did little to illuminate social motivations and pragmatic means and functions of English-sourced swear words compared to heritage Finnish swear words. With phase 2, we built on the questions that were raised as a result of phase 1 or that remained unaddressed after phase 1. Specifically, with this portion of the study we explore the following research questions:

RQ1) What are the social and stylistic associations of the borrowing *shit* compared to the heritage Finnish form *paska*?

RQ2) What are the social characteristics associated with the English loan word *fuck*? How pragmatically loaded is it compared to the Finnish heritage form *vittu*?

RQ3) What are the social and stylistic associates of the borrowing *damn* compared to the heritage Finnish forms *hitto* and *perkele*?

##### 4.1 The online survey data

The data for this portion of the study was collected through an online survey that made use of the classical matched guise techniques (Lambert 1967). The survey was created using QuestionPro software. The survey consisted of a biographical section, four sections asking respondents to assess audio samples containing target swear words, and a final section asking respondents to assess and self-report on three written sentences in Finnish containing the swear words *hitto*, *damn* and *perkele*. The survey was piloted with two target groups before going live in March 2018, and it was tested with potential users on different kinds of electronic devices. During a five-week data collection period, there were 446 finished responses to the survey. While the number of respondents was more modest than expected (with a relatively high interruption rate perhaps due to constraints of listening to audio samples), the respondents were nonetheless relatively optimally divided across age groups, education levels and regions of Finland. The breadth of distribution was the result of using multiple channels to spread the survey link, for example, collaboration with the national Community College network and online groups of relatively older portions of the population. The gender distribution was 46 percent male, 48 percent female and 6 percent other/not applicable. There was slightly heavier representation among the more urbanized southern portion of Finland, particularly the Helsinki capital region, and among people who have at least some tertiary education.

The audio portion of the matched guise test made use of recorded utterances from three native speakers of Finnish. The voices were selected, piloted and ultimately represented on the survey as examples of a middle-aged man, a middle-aged woman and a teen-aged girl. In this task set, respondents to the survey heard altogether 13 audio samples consisting of three different swear word/expletive variables: *shit/paska*, *oh my god* and *fuck/vittu*.

#### 4.2. Results of online survey

##### *RQ1) shit compared to paska*

The aim of this portion of the survey was to assess the geographical dialectal perceptions of the borrowing *shit* compared to the heritage form *paska* and also their relation to standardized Finnish. To achieve this aim, we made use of six recorded utterances from the same middle-aged male voice. We used just one voice for these utterances to reduce language-external variation; gender and age were not pertinent features for this portion of the study (see Vaattovaara and Peterson 2019 for further details). The base sample sentence was, in standardized Finnish, *ihan samaa vanhaa paskaa* ‘just the same old shit.’ The audio samples feature six variations of this base sentence. The six audio samples featured variation in regional and standard/non-standard pronunciation, as well as in the use of either *paska* or *shit*. We coached the speaker to utter the same sentence in six different styles, which are summarized in Table 3: standardized Finnish, urban Finnish (associated with the Helsinki region), and a style exhibiting two non-standard but widespread dialect features (namely, nonstandard gemination and epenthesis).

‘just the same old shit’	phonological style, EN/FI variant	order in survey
1) <i>ihan samaa vanhaa paskaa</i>	standard, FI	1/11
2) <i>ihan samaa vanhaa shittij</i>	urban, EN	4/11
3) <i>ihan samaa vanhaa shittiä</i>	standard, EN	5/11
4) <i>ihan sammaa vanhaa shittiä</i>	rural+standard, EN	7/11

5) <i>ihan sammaa vanahaa paskaa</i>	rural, FI	10/11
6) <i>ihan sammaa vanahaa shittii</i>	rural+urban, EN	11/11

Table 3. Audio samples with variable *paska* and *shit*, produced by a middle-aged male (from Vaattovaara and Peterson 2019)

Table 3 shows the phonological variation that was tested in this portion of the survey. The main variants, as discussed previously, are the variable *shit/paska*. The phonological variations explored are, in order:

- 1) standard Finnish utterance with variant *paska*, here in the partitive form
- 2) urban (Helsinki region) phonological form with *shittii*, exhibiting vowel assimilation of the partitive ending *-ii*. The standard partitive ending is *shittiä*, as exhibited in samples 3) and 4). This subtle but salient phonological difference is decidedly associated with southern and urban styles of spoken Finnish.
- 3) standard Finnish utterance with variant *shit*, here in the (standard) partitive form
- 4) a combination of non-southern phonology, exhibited with the nonstandard geminate in *sammaa* ‘same’ and nonstandard epenthesis in *vanahaa* ‘old,’ co-occurring with the standard partitive form of *shit*.
- 5) like 4), a combination of non-southern, nonstandard phonology (gemination and epenthesis), coupled with the standard partitive form of *paska*.
- 6) a combination of non-southern features, as exhibited in samples 4 and 5, co-occurring with the nonstandard, urban southern partitive form of *shit*.

For each utterance, respondents were asked to assess on a 7-point Likert scale how “available” the utterance was and how “acceptable” it was. These Likert scale assessments were analyzed using linear mixed models, which allowed us to distinguish the population level effects from person-wise effects. The results of this portion of the survey are described in detail in Vaattovaara and Peterson (2019). In summary, the results showed that, as expected, the audio samples containing the heritage variant *paska* were deemed as most “available,” with the standard Finnish utterance, as shown in 1), receiving the highest ranking. Utterance 5), featuring non-southern phonology and the heritage form *paska*, was ranked second. The least “available” utterance was audio sample 4), which was apparently perceived as being a mismatch of southern urban and non-southern styles. There was no significant difference in the acceptance of this utterance and audio sample 6). The results are shown in Table 4.

voice sample	availability (mean rate)	StD
Ihan samaa vanhaa paskaa (standard)	2.39	1.43
Ihan sammaa vanahaa paskaa (rural)	3.66	1.91
Ihan samaa vanhaa shittii (urban)	4.12	1.87
Ihan samaa vanhaa shittiä (standard)	4.75	1.82

Ihan sammaa vanahaa shittii (non-southern + urban)	5.19	1.77
Ihan sammaa vanahaa shittiä (non-southern+standard)	5.37	1.8

Table 4: “Availability” of sample phrases, with 1 = most available, 7 = least available (adapted from Vaattovaara and Peterson 2019)

In terms of “acceptability,” the other value offered on a Likert scale, assessments clustered closely together (4.98–5.09) without significant differences (for further details see Vaattovaara and Peterson 2019: 4).

We interpret these results as indicating, in line with our previous studies on pragmatic borrowings from English, an association of the borrowed forms with the southern and urban region of Finland. This is evident, for example, from the results regarding sample phrase 2), the southern and urban variant of the partitive form of *shit*, which ranked highest among the four *shit* sentences in terms of availability. In terms of the “acceptability” scale, there were very small scalar distinctions between the *paska* and *shit* audio samples and no statistical differences. The similarity in the reaction to the two forms points toward an overall lack of distinction in pragmatic force and meaning between the heritage and the borrowed variants.

#### RQ2) fuck compared to vittu

In the survey, respondents heard the same middle-aged male voice utter a phrase that differed with regard to a Finnish heritage exclamation versus a borrowed English exclamation. These two phrases were *siis mitä vittua* and *siis what the fuck*. The discourse particle *siis* ‘well’ was included as the initial element of both phrases to signal to the listeners that the matrix language was Finnish. In addition, the speaker of the English-sourced phrase was instructed to utter the phrase with a Finnish intonation pattern and pronunciation. After hearing each phrase, respondents were presented with a scale labeled with the statement “The sentence is in my opinion ordinary/natural Finnish,” and six options labeled “completely agree,” “slightly agree,” “neither agree nor disagree,” “completely disagree” and “I don’t know.”

Not surprisingly, a majority of the respondents, 89 percent, considered the phrase *siis mitä vittua* to be at least somewhat typical or natural in Finnish. None of the respondents chose the “I don’t know” option. For the phrase *siis what the fuck*, only 26 percent of the respondents reported that it was either slightly or entirely ordinary/natural Finnish. The majority, 65 percent of the respondents, considered the English borrowing unnatural, although, as shown in Figure 2, there were fewer respondents who regarded the phrase as completely unnatural compared to those who found the phrase only partially unnatural. This finding indicates at least some level of integration into Finnish, even with a grudging acceptance of the form.

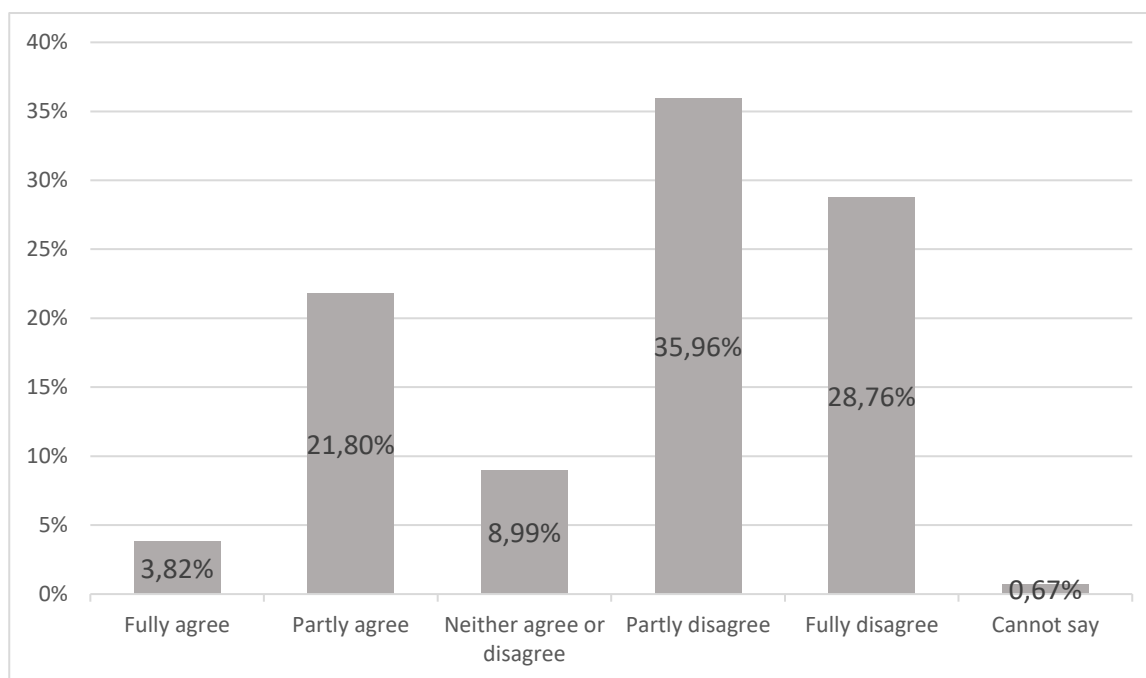


Figure 2: Responses to the audio sample *Siis what the fuck*: Is it ordinary/natural Finnish?

In terms of the linear regression analysis, the results showed that the phrase containing the borrowing *fuck* was statistically less available to respondents over the age of 60. Age was the only significant social factor. Insignificant factors were gender, level of education, size of municipality and geographical region.

The oral matched guise survey did not ask respondents to assess the pragmatic weight of the two phrases respective of each other. To gain understanding of how potentially offensive *fuck* is in Finnish compared to the heritage form *vittu*, the open-ended responses were helpful. Out of 446 respondents, 168 took the opportunity to report on their attitudes toward these forms (see examples 9 and 10).

(9) *Vittua kuulee joka paikassa - siis aivan tavallista suomea mutta ei oikein hyväksyttävää kielenkäyttöä muodollisissa yhteyksissä.* ‘*Vittu* is heard everywhere – I mean it is totally normal Finnish language but not really acceptable language use in formal situations.’

(10) *ilmaisuu on nuorten puheessa yleinen, mutta mielestäni ei tavallista eikä luontevaa suomea* ‘the phrase is common in youth speech, but in my opinion it is not typical or natural Finnish’

These examples as well as many others demonstrate that, while there is acknowledgement that *vittu* is a swear word, it is nonetheless an everyday facet of Finnish discourse. Open-ended responses about the phrase containing the borrowing *fuck* were more divisive. While some respondents criticized the phrase for containing an English borrowing, several commented on the reduced pragmatic weight of the borrowing compared to the heritage Finnish form.

(11) *Vieraalla kielellä kiroilu on laimeampaa, siinä ei ole samaa alkuvoimaa kuin äidinkielen manauksissa.* ‘Swearing in a foreign language is milder, it doesn’t have the same sense of primal power as the swear words of one’s first language.’

The opinion expressed in (11) is reminiscent of the body of work carried out by Dewaele (for example, Dewaele 2004, 2010), which routinely points out that swear words in another language tend to have a lower pragmatic force than in the mother tongue(s). An interesting addition to these findings is that a swear word borrowing into a recipient language apparently has a lower pragmatic force, as well.

### RQ3) *damn* compared to *perkele* and *hitto*

Phase 1, the corpus study, demonstrated that *damn* is one English-sourced swear word that is used as a particle in Finnish discourse. With this finding in mind, we wished to gain further insight into its use in Finnish, specifically in its particle function. To this end, we tested the borrowing *damn* in an utterance gained from the corpus study. The sample utterance, taken from the Suomi24 corpus, appeared in our online survey not as part of the oral matched guise activity (i.e, not as an audio sample), but rather as a set of written sentences. Survey respondents were instructed to first read the following sentences:

*Ai, olihan siellä, hitto, en heti nähnyt.*  
*Ai, olihan siellä, damn, en heti nähnyt.*  
*Ai, olihan siellä, perkele, en heti nähnyt.*

(‘Ah, it was there, ----- . I didn’t see it right way.’)

Beneath the three sentences respondents were asked to answer the question “Which of these would you most commonly/probably use yourself?” They were then given the option to choose any number of the following options: *hitto*, *damn*, *perkele*, or *I wouldn’t use any of these*. Finally, respondents were invited to write their reactions to these forms in a free-form text box with the request “Why would you use this/these ones? What else would you like to say about the swearwords *hitto*, *damn*, *perkele*?”

The reason for testing this particular semantic set was, in particular, to try to assess the pragmatic weight of *damn* compared to the heritage forms *hitto*, a mild variant. *Perkele*, as described earlier, is considered one of the oldest and most authentically Finnish of all heritage swear words (Hjort 2014, 2017). We also wished to tease out any perceptions of the relationship of these variants in relation with social categories such as age, region and gender.

Of the three choices, *hitto* by far received the highest number of claimed ownership among respondents. Out of the total 446 respondents, 264 claimed they would use *hitto*, 151 claimed they would use *perkele*, and nearly even levels of respondents claimed they would use *damn* or *none of these*: 66 and 49, respectively.

The results of our linear mixed model showed that the heritage forms *perkele* and *hitto* are reported by the respondents as being more acceptable than *damn*. With *damn* as the reference category, regression analysis showed that age is the best predictor for the users of *hitto*, the mild variant, with respondents over age 60 showing a preference for this form ( $p=0.037$ ). Age was the only significant factor for favoring *hitto* over *damn* and *perkele*. Non-significant factors included gender, level of education, geographical region and size of municipality. An interesting outcome from this particular statistic model is the amount of indecision apparent in the responses. With a 95 percent confidence interval [0.216; 6.973] and a coefficient of 3.6, the results show that there is not much agreement among respondents as to the acceptability of these three variants relative to each other.

There was an interesting binary gender<sup>3</sup> split in the data in two respects: men tended to claim to be the users of the strong form *perkele*, whereas women claimed to be the users of *damn* ( $p=0.0169$ ). While based on previous studies (e.g., Hjort 2017) we expected level of education to play a role in determining the use of *perkele*, education turned out to be not significant. Other non-significant factors included geographical location and size of municipality.

More than any other of the swear word sets we observed, gender came to the forefront with this particular semantic set. This was the only set for which gender was a significant factor in the linear mixed model, and this was also the set which elicited the most open-ended written comments relating to gender. Sample responses are shown in (12) and (13).

12) *Käytän hitto ja damn. En käytä perkelettä, koska sitä käyttää vanhat lande sedät :) 'I use hitto and damn. I don't use perkele, because it is used by old guys from the countryside :)'*

13) *Olen mies, hitto sopii paremmin naisen suuhun ja damn jättää miehestä epämiehekkään/vähä-älyisen vaikutelman. 'I'm a man, hitto is better fit for a woman's mouth and damn gives a non-masculine/idiotic impression of a male'*

With regard to the question of pragmatic weight, the open-ended responses were more illuminating than the statistical analysis. Of the 393 open-ended responses received, a large portion of them commented on the perceived pragmatic equivalence of the mild heritage variant *hitto* and *damn*. At the same time many comments were critical of the use of an “English” word in the middle of Finnish discourse in a first place.

## 5. Conclusions

Our study, based on two-phased analysis, revealed several interesting yet unexpected findings about the use of swear words in Finnish. First, the corpus phase of the study indicated that heritage Finnish swear words are not being displaced by English swear words. Rather, English-sourced swear words, as shown in our previous studies of English pragmatic borrowings in Finnish, are relegated to a specific set of social and pragmatic circumstances, and they are often viewed critically. Second, the corpus phase of the study showed that, contrary to our expectations, swear-word borrowings from English are not favored for use in discourse-particle functions. Rather, although in general English swear words appear to be dispreferred, they are nonetheless seamlessly integrated into Finnish grammar, which hints at long-term exposure and also a lack of markedness. In the corpus data explored, which originates from a Finnish online forum, the most frequent English-sourced swear words also featured prominently in creative, apparently performative code-switches.

Of particular interest are the outcomes related to each semantic set of swear words we investigated, or in other words the heritage / borrowed sets, *shit/paska*, *fuck/vittu* and *damn/perkele/hitto*. The outcome of phase 2 of our study, an online survey with a modified matched guise test and a self-report task, demonstrated distinct perceptions and social values associated with each separate borrowing from English. This, in line with the outcome of phase 1, points toward a high level of integration of these borrowed forms. The most important outcome regarding the *shit/paska* pair was the reluctance of survey respondents to accept the borrowing *shit* with non-southern Finland pronunciation features, highlighting that this borrowing—and, presumably, other

---

<sup>3</sup> Other genders were not considered in the analysis, as there were too few participants of nonbinary genders to be tested through our statistical model.

borrowings from English—are associated with the more urbanized southern part of Finland or the perceived “southern urban speech style”. The findings regarding the preferred pronunciation of *shit* in the partitive form underlined this preference: the form was preferred in the partitive variant *shittii* (instead of the standard *shittiä*), which is highly correlated with the urban Helsinki area.

The joint findings from phases 1 and 2 of the study highlight that the borrowing *fuck* is indeed an integrated component of contemporary Finnish discourse. In our data, this form was manifest in an almost surprising array of Finnish morphological structures, word formations and bilingual puns. Phase 2 of our study revealed age to be the single most important factor in determining the availability of *fuck*: those over age 60 claim it is not available (or in other words, not common) in Finnish, while those under age 60 claim that it is.

For the most part, among the swear word sets investigated, gender did not emerge as a significant factor. In light of previous claims on the relationship of gender and swearing, this outcome is quite satisfying: for all of the semantic sets we looked at, there was no significant difference in the claimed behavior of women compared to men. The exception was the *damn/hitto/perkele* set. The most compelling outcome of this portion of the study is that gender appears to be the determining factor for *perkele* versus *damn*, at least based on these data. *Perkele* retains its status as the quintessential Finnish swear word, associated, as revealed through the open-ended responses to our survey, with non-urban men. *Damn*, on the other hand, is more likely to be associated with women, and, as such appears to offer women an innovative means of expressing the same or at least overlapping pragmatic functions, but without using a swear word associated with masculinity.

## References

- Aller Media ltd. 2014. *The Suomi 24 Sentences Corpus (2016H2)* [text corpus]. Kielipankki. Retrieved from <http://urn.fi/urn:nbn:fi:lb-2017021505>
- Aller Media ltd. 2019. Suomi24-tilasto. Retrieved from [https://www.aller.fi/wpcontent/uploads/2019/01/S24\\_Fiam\\_2018.pdf](https://www.aller.fi/wpcontent/uploads/2019/01/S24_Fiam_2018.pdf)
- Andersen, Gisle. 2014. Pragmatic borrowing. *Journal of Pragmatics* 67: 17–33. 10.1016/j.pragma.2014.03.005
- Beers Fägersten, Kristy. 2014. The use of English swear words in Swedish media. In *Swearing in the Nordic Countries*, edited by Marianne Rathje, 63–82. Copenhagen: Dansk Sprognævn.
- Beers Fägersten, Kristy. 2017. The role of swearing in creating an online persona: The case of YouTuber PewDiePie. *Discourse, Context & Media* 18: 1–10. <http://doi.org/10.1016/j.dcm.2017.04.002>
- Bloomfield, Leonard. 1933. *Language*. New York City, NY: Holt.
- Borin, Lars, Markus Forsberg, and Johan Roxendal. 2012. *Korp – the corpus infrastructure of Språkbanken*. Proceedings of LREC 2012. Istanbul: ELRA: 474–478.
- Dewaele, Jean-Marc. 2004. The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development*, 25 (2–3): 204–222. <https://doi.org/10.1080/01434630408666529>
- Dewaele, Jean-Marc. 2010. “Christ fucking shit merde!” Language preferences for swearing among maximally proficient multilinguals. *Sociolinguistic Studies*, 4 (3): 595–614.
- Hjort, Minna. 2014. *Perkele* - a literary companion: A glimpse at the function and frequency of the Finnish swearword in translated and non-translated contemporary fiction. In *Swearing in the Nordic Countries*, edited by Marianne Rathje, 125–149. Copenhagen: Dansk Sprognævn.

- Hjort, Minna. 2017. Swearing in Finnish: Folk definitions and perceptions. In *Advances in swearing research: new languages and new contexts*, edited by Kristy Fägersten and Karyn Stapleton, 231–256. Amsterdam: John Benjamins.
- Kulonen, Ulla-Maija. 1990. Miten suomalaiset kiroilivat ennen kristinuskkoa? *Kielikello* 1/1990, 3–6.
- Lagus, Krista, Mika Pantzar, Minna Ruckenstein and Marjoriikka Ylisiura. 2016. *Suomi24: Muodonantoa aineistolle*. Kuluttajatutkimuskeskus. Valtiotieteellisen tiedekunnan julkaisuja 2016:10. Helsinki: Helsingin yliopisto.
- Lambert, Wallace E. 1967. A social psychology of bilingualism. *Journal of Social Issues* 23 (2): 91–109.
- McEnery, Tony. 2006. *Swearing in English: Bad Language, Purity and Power From 1586 to the Present*. Routledge Advances in Corpus Linguistics. London: Routledge.
- Müller, Simone. 2005. *Discourse markers in native and non-native English discourse*. Amsterdam: John Benjamins.
- Peterson, Elizabeth. 2017. The nativization of pragmatic borrowings in remote language contact situations. In *The Pragmatics of Borrowing*, edited by Cristiano Furiassi, Gisle Andersen and Biljana Mišić Ilić. *Journal of Pragmatics* 113: 116–126.  
<https://doi.org/10.1016/j.pragma.2017.02.012>
- Peterson, Elizabeth and Johanna Vaattovaara. 2014. *Kiitos* and *pliis*: the relationship of native and borrowed politeness markers in Finnish. *Journal of Politeness Research* 10 (2): 247–269.  
<https://doi.org/10.1515/pr-2014-0011>
- Pyykkö, Katariina. 2017. Borrowing *jees* as a Finnish adjective: A case of nativization. Unpublished MA thesis, Department of Languages, University of Helsinki.
- Stapleton, Karyn. 2003. Gender and Swearing: A Community Practice. *Women and Language* 26 (2): 22–33.
- Stapleton, Karyn. 2010. Swearing. In *Interpersonal Pragmatics*, edited by M.A. Locher and S. Graham, 289–305. Berlin: De Gruyter Mouton.
- Vaattovaara, Johanna and Elizabeth Peterson. 2019. “Same Old Paska or New Shit? On the Stylistic Boundaries and Social Meaning Potentials of Swearing Loanwords in Finnish.” *Ampersand* 6: 100057. <https://doi.org/10.1016/j.amper.2019.100057>.

## Glossing conventions

- 1SG – first-person singular  
 2SG – second-person singular  
 3SG – third-person singular  
 ADE – adessive  
 CLITIC – clitic  
 COND – conditional  
 GEN – genitive  
 ILL – illative  
 ILL – illative  
 INANIM – inanimate  
 INF – infinitive  
 INF.III – “3<sup>rd</sup> infinitive”, a nominalizing infinitive illative  
 INF.IV – “4<sup>th</sup> infinitive”, a nominalizing infinitive  
 NOM – nominative  
 PL – plural  
 PTV – partitive