

Codeswitching between English and French on an online discussion forum

Margus Kirsimägi

Master's Thesis

Language Specialist Degree Programme, Department of English

School of Languages and Translation Studies

Faculty of Humanities

University of Turku

May 2024

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service.

Master's Thesis

Language Specialist Degree Programme, Department of English

Margus Kirsimägi

Codeswitching between English and French on an online discussion forum

Number of pages: 46 pages, 7 pp. appendices

This master's thesis examines codeswitching between English and French on an online discussion forum. The purpose of the study is to examine in what contexts does codeswitching take place, what types of codeswitching occur in the discussion forum, and what are the functions of different codeswitches.

The materials consisted of 582 forum posts that were analysed for French codeswitches. 150 posts in total contained codeswitching. The analysed codeswitches were divided into intersentential, intrasentential or metalinguistic commentary categories on the basis of the structure of codeswitches. Intrasentential switches are also categorised into nouns, adjectives, verbs etc. The codeswitches were also analysed in terms of their function in a context and divided into several different categories, such as cultural switches, language discussion, quotations, and humour.

The results show that intrasentential codeswitching occurs much more frequently than intersentential switching. Intrasentential switches consisted of different French words or short French phrases. The functional analysis showed that cultural switches occur much more frequently than any other function category in the data. Cultural switches often referred to objects and concepts specific to French culture and society. *Language discussion* codeswitches were also quite frequent. These involved discussion about the French language, how to translate certain words into English or discussion about the meanings of different French words.

Key words: Codeswitching, French, English, Computer-mediated communication

Table of contents

1	Introduction	1
2	Background	3
2.1	Codeswitching	3
2.2	Codeswitching vs borrowing	6
2.3	Functions of codeswitching	7
2.4	Web forum language	8
3	Previous research on codeswitching in electronic writing	10
4	Research methodology	15
4.1	Materials	15
4.2	Methods	20
5	Analysis	22
5.1	Structural analysis	22
5.1.1	Nouns	23
5.1.2	Noun phrases	26
5.1.3	Proper nouns	27
5.1.4	Other categories	27
5.2	Functional analysis	29
5.2.1	Cultural switches	30
5.2.2	Free switching	32
5.2.3	Language discussion	33
5.2.4	Quotations	34
5.2.5	Humour	35
5.2.6	Greetings/farewells	36
5.3	Discussion	37
6	Conclusion	41
	References	44
	Appendices	47
	Finnish Summary	47

1 Introduction

This Master's Thesis focuses on codeswitching between English and French on the online discussion forum *CompleteFrance*, which is a forum meant for discussing anything about France, such as culture, finance, moving, language, or literature. Codeswitching takes places when more than one language is used in spoken or written context (Myers-Scotton 2009, 4), but it can also mean the use of two or more linguistic varieties or dialects in a conversation (Gardner-Chloros 2009). The meaning of the term has some variation as it has been defined slightly differently at different times by various researchers. It depends on what kind of codeswitching the researchers are investigating and how they approach the topic.

Both English and French are widely spoken across multiple continents, although English has notably more speakers than French. In addition, English serves as a *lingua franca* for a considerable number of people. It is used in many different academic and professional circles as it is a language of science, technology, media, and globalisation (Kuzelewska 2020, 1424). Furthermore, another research article points out that “no other language has occupied the position English holds today” (Barančicová and Zerzová 2018, 30) since no other language has so many speakers. According to the WorldData (2023), there are about 272 million native speakers of English and about 411 million people speak English as their mother tongue. In addition, “English is an official language in 39 countries and is partly spoken as a mother tongue in 19 other countries” (ibid.).

There are many different discussion forums for all sorts of different topics on the internet, and they are important communication environments nowadays. Each forum has its own community and a purpose, and discussion forums offer plenty of linguistic data for analysing different topics, such as multilingual practices, language usage or codeswitching. Analysing the language of different discussion forums can reveal how members of the forum communicate with each other, how different users employ language and what language norms or rules exist on the forum. It is also possible to analyse how members signal their membership and construct identity by using multilingual practices (Kortelainen and Kolehmainen 2022, 8). Discussion forums can use only one language which can act as the official language, but they may also be multilingual where users use two or even more languages frequently on the forum.

The purpose of this thesis is to investigate the contexts in which codeswitching between English and French takes place, and what the functions of different codeswitches are in the discussion forum. The research questions of this study are: In what contexts do people switch most often in a discussion forum? What types of codeswitching occur in the discussion forum? What are the functions of switches? This topic is relevant as there are not many studies about codeswitching between English and French. I would like to examine the structure of codeswitches, and what functions codeswitches have in written texts. English and French specifically are a good pair for this as English is a global language and is essential for communication in various fields over the world while French is also a widely spoken language.

2 Background

In this chapter, I introduce the meaning of codeswitching, key terms and how they are defined in 2.1. Codeswitching association with borrowing is investigated in 2.2 and the functions of codeswitching are discussed in 2.3.

2.1 Codeswitching

Codeswitching has received multiple slightly differing definitions and competing terms over the course of time, which can depend on what aspect is under study and by what methods is the topic approached. The term often refers to “the use of more than one language in the course of a single communicative episode” (Eastman and Heller 1988,1). This communicative episode can be spoken or written. Codeswitching can also take place when using only one language, but one shifts between different varieties or dialects (Garner-Chloros 2009, 4). Codeswitching affects everyone at least in some extent who are in contact with two or more languages or dialects (ibid.).

Researchers tend to distinguish between structurally and functionally different multilingual practices as some researchers distinct between intersentential codeswitching and intrasentential codeswitching for instance (Kortelainen and Kolehmainen 2022, 11–12). Codeswitching between sentences or clauses is called intersentential codeswitching while codeswitching within a sentence or clause boundaries is referred to as intrasentential codeswitching (Myers–Scotton 1993, 3–4). Intrasentential switching has been studied more since it offers more structural topics since it involves two grammars coming into contact with each other while in intersentential switching they do not and the codes remain structurally separate (Kortelainen and Kolehmainen 2022, 12). Next, I will discuss *conversational* and *situational* codeswitching.

According to Gumperz (1982, 59), *conversational* or *metaphorical* codeswitching has interactional meaning and it “[...] can be defined as the juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or subsystems”. In most cases, the alternation takes the form of two subsequent sentences when a speaker either repeats what someone said or to reply to a statement in a second language (ibid.). Rajala (2019, 9) notes that “[i]n conversational code-switching, there is no one-to-one relationship between extralinguistic context and language use in conversation”. Due to that, it is possible to utilise conversational codeswitching metaphorically for separating old

information from new, “marking the degrees of emphasis or contrastiveness”, keeping the topic and subject separate or signalling the speaker’s position towards the message they produced (Gumperz 1982, 48). According to Rajala (2019, 9–10), “Therefore, speakers (or writers) do not respond to a fixed social context but rely on their own and their interlocutor’s abstract understanding of situational norms to communicate meaning through codeswitching”. The relation between group identity and code-choice is in that case a symbolic one and for that reason, “code-switching is used as a metaphorical tool that conveys meanings attached to the different codes: i.e., code-switching acts as a metaphor for a socially motivated message” (ibid.).

Gumperz (1982, 60) defines *situational* codeswitching in the following way: “[d]istinct varieties are employed in certain settings (such as home, school, work) that are associated with separate, bounded kinds of activities (public speaking, formal negotiations [...]) or spoken with different categories of speakers (friends, family members, strangers [...])”. These settings, activities and speakers belong to conversation-external factors. For example, in the old Catholic Latin phrases were read among other local languages and in some tribal societies some parts of public addresses were translated and repeated in other languages as it was required by etiquette (Gumperz 1982, 60–61). The etiquette rules that govern language use are explicitly taught in most cases and breaches can evoke overt comment as the norms of code selection tend to be stable (ibid.) According to Gumperz (1982, 61), “[y]et the alternation always corresponds to structurally identifiable stages or episodes of a speech event”. “Both conversationalists and linguists agree in assigning each sentence or group of sentences to one code or another” (ibid.). There is almost one to one relationship between language use and social context and each variety can have a distinct place and function in the local speech repertoire (ibid.). Language communities tend to use different languages or varieties depending on the setting. Some languages might be more favoured in certain contexts depending on the speakers, location, and topic. For example, in certain parts of Canada, French holds significantly more prestige than English, so in many circles French is the preferred language of communication, but it is not expected that all English speakers speak French fluently. Similarly, in France particularly older people might not have good English skills and language attitudes might also play a role in choosing the language for a setting. Religion, age, gender and class factors also play a role in deciding which languages are used.

One often used theory of codeswitching in research is the Matrix Language Frame mode theory, which was developed by Myers-Scotton in 1993, which explains what is *classic codeswitching*. According to Myers-Scotton (2011, 241–242) “[c]lassic codeswitching includes elements from two (or more) language varieties in the same clause, but only one of these varieties is the source of the morphosyntactic frame for the clause”. In intrasentential codeswitching, the language that supplies the morphosyntactic frame in codeswitched discourse is called the Matrix language or dominant language while the less dominant language is called the Embedded Language (Myers-Scotton 1993,3–4). In this study, English acts as the dominant language which supplies the morphosyntactic frame and elements from French, which is the Embedded Language, are placed into this frame.

Kortelainen and Kolehmainen (2022, 12) note that “[i]n codeswitching research, one major branch has focused on the syntactic properties and grammatical integration of elements in codeswitched discourse”. For instance, it can be studied how elements of grammar of the embedded language or other multilingual elements are structurally integrated in sentences where codeswitching takes place. French grammar works differently in several aspects compared to English. For instance, English and French use articles differently. All French nouns are either masculine (*un* for indefinite and *la* for definite) or feminine (*une/la*) and the article must match the gender of its noun while English nouns have no grammatical gender. Thus, in codeswitching it is possible that English nouns have a French article. Similarly, most of the adjectives in French come after the noun, not before (*une maison élégante*, ‘an elegant home’). English and French have many similar looking words which can have the same meaning or a different one (for example, coin vs *coin*. *Coin* in French refers to a ‘corner’. *Coin* in context of money is *une pièce de monnaie*). Another aspect to keep in mind is that the use of tenses differs between the languages (particularly, when referring to events that have happened in the past). Capitalisation rules are also different as in French languages are generally not capitalised. In titles, only proper names and the first letter of the title are capitalised. In terms of punctuation, French language adds spaces after many punctuation marks (question mark, exclamation point etc), which is a very visible feature in the discussion forum I analysed. Furthermore, French language uses different quotation marks (« ») than English.

In terms of functional multilingual practices, Auer (1999, 310) regards the term codeswitching for cases where individual switches carry communicative meaning and are perceived as locally meaningful by the participants. In this study, I use the term codeswitching in its general sense, meaning when it refers to the alternating use of different languages, and there are only two studied in this thesis.

2.2 Codeswitching vs borrowing

Codeswitching and borrowing share some characteristics. The distinction between a single word being a single-word codeswitch or a borrowed lexical item can be a challenge, which is discussed by many researchers, such as Rajala (2019) and Kortelainen and Kolehmainen (2022) in their research. Gumperz similarly (1982, 66) notes that “code switching must be separated from loan word usage or borrowings” and he continues by stating that “[b]orrowing can be defined as the introduction of single words or short, frozen, idiomatic phrases from one variety into the other”. These words or phrases become part of the grammatical system of the language that borrowed them and “[t]hey are treated as part of its lexicon, take on its morphological characteristics and enter into its syntactic structures” (ibid.).

Codeswitching “[...] relies on the meaningful juxtaposition of what speakers must consciously or subconsciously process as strings formed according to the internal rules of *two distinct grammatical systems*” (Gumperz 1982, 60). Single words in the embedded language can be considered as codeswitches because elements from single words to full phrases in bilingual clauses fall on a continuum of embedded language (Myers-Scotton 2011, 253-260). Rajala (ibid.) also considers words included in the *Oxford English Dictionary* to be borrowings and not codeswitches. I will follow the same criterion in regard to single-word codeswitches in the materials, unless the OED specifically states that the word is specific to France or it is obsolete in English. Furthermore, some words can have different meanings in English and French. If the French meaning is used, then I consider it to be a codeswitch. However, it should be noted that for Rajala’s study, the distinction between a codeswitch and a borrowing is not important since his research questions and approach are different, but it is relevant for my study because I focus specifically on different kinds of codeswitching and their functions. Mostly codeswitching utilises words from another language that are likely not very often used by the speakers or writers as they are not part of the vocabulary of the dominant language.

2.3 Functions of codeswitching

The functions of codeswitching have also been studied extensively. Codeswitching can be used to display identity or to achieve different communicative goals, and of significance is also the sequential position of codeswitches within a discourse and explaining what interactional purpose they have at any point, which has been studied for example by Auer (1984) and Wei (1998, 280–316).

Codeswitching functions as a *contextualisation cue*, which are “means by which speakers signal and listeners interpret what the activity is, how semantic content is to be understood and how each sentence relates to what precedes or follows” (Gumperz 1982, 131). These cues are often used and perceived by people, but we almost never talk about them directly and only notice them unconsciously in most cases (ibid.). Codeswitching as a contextualisation cue works in a similar way as posture, intonation, or tone as it helps the speaker to indicate to other participants how they intend what is said and how they wish their content of message to be understood by other participants (Gumperz 1982, 131; Kortelainen and Kolehmainen 2022, 13). Contextualisation cues carry information, but their meanings are implicit as they are not discussed out of context, but instead they are conveyed as part of the interaction process (Gumperz 1982, 131). Typically, participants notice and understand necessary cues and the interpretive process is taken for granted and goes unnoticed, but sometimes it is possible that a listener does not notice a cue or does not understand its function, or interpretations differ, in which case a misunderstanding occurs (Gumperz 1982, 132). In that case if the listener is notified of this difference, it tends to be seen in attitudinal terms and the speaker is said to be rude or uncooperative (ibid.). However, participants do not usually notice that the listener may have failed to notice a change in pronunciation or rhythm and miscommunication of this type is regarded as a social indiscretion and leads to misinterpretation of the speaker’s intent and it is not likely to be identified as a linguistic error (ibid.).

Codeswitching is also related to identity which is often demonstrated by terms: we-code and they-code (Gumperz 1982). The language used by the minority group acts as a we-code and is for in-group and informal activities while they-code is the language of the majority and is associated with formal, out-group activities (Gardner-Chloros 2009, 56). Although it should be noted that the relationship “between communicative style and group identity is a symbolic one” since it is not possible to measure actual usage and there is no direct connection “between the occurrence of a particular set of linguistic forms and extralinguistic context”

necessarily and only in certain rare contexts is only one code appropriate (Gumperz 1982, 66). Due to that, there are empirical difficulties in describing each person's perceptions of what counts as we-code and they-code (ibid.). It is similarly noted by Sebba (1998) that the status of we-code and they-code is not stable, and according to Bailey (2007), identities and identity associations of codes are built in the course of an interaction instead of being predetermined. Different identities can be emphasised at different points during one interaction and one code can take the role of either we-code or they-code depending on the context (Kortelainen and Kolehmainen 2022, 14). They argue that "[t]here is no one-to-one match between specific codeswitching functions and identity presentation but the interpretation of an individual switch as an enactment of identity is rather done on the basis of the entire discourse and its larger context" (ibid.). In my study, I also explore the role of English and French and to which code they belong in specific contexts. There are some topics in the forum where discussion takes place entirely in French and in some topics random posts are written mostly in French, but other people answer in English. It is likely that in these cases, the person does not know English well or refuses to use it.

2.4 Web forum language

Studying codeswitching in writing has various challenges. One reason is that there are no frameworks that are specifically designed for studying codeswitching and language choice in digital media (Lexander 2012, 146). Lexander (2012, 146) notes that many researchers have applied various theories that were originally meant for studying spoken language, such as Myers-Scotton's markedness model (Myers-Scotton 1998, 18–38) and Gumperz' situational and metaphorical codeswitching (Gumperz 1982, 55–99). Other theories include social network analysis (Paolillo 1999) and Peter Auer's conversational approach (1995, 1998). Some researchers have chosen an eclectic approach, like Lars Hinrichs (2006), who draws partly on the markedness model and also on Auer's theory and Gumperz (1982). However, these theories do not fully account for the multilingualism of written texts since they were developed in regard to spoken language research (Hinrichs 2006), but they are still worth noting here as they offer useful insight into studying codeswitching.

Furthermore, while there are many similarities between interactive web writing and oral conversation, the language that is used in web forums is seen as a distinct form of writing that also has multiple visual means and the aural modality is mostly absent (Kytölä 2012, 109). There is no sound typically in web forums (although it is possible that some forums contain

audio files, which are part of the discussion in the topic thread), and for that reason language used in web forums is not considered to be spoken language (Kytölä 2012, 109). According to Kytölä (2012, 109) “[o]ral communication can be spontaneous, unplanned, and unconscious with regard to multilingualism, but it can equally well be planned, ‘performed’ and consciously stylized”. It depends on the situation, participants and genre instead of modality (ibid.). The same applies to web forum language as it is written and asynchronous, “yet it can range from spontaneous, unplanned or unconscious to highly stylized, carefully constructed, even rhetorical essays” (ibid.). Members of the forum decide what kind of language they are writing and adapt the forum to their specific needs (ibid.). For example, one can decide to carefully construct their message in a discussion forum by using sources and dictionaries while another person can write fast without using any other sources (ibid.). These patterns can result in very different multilingual outputs in terms of language choice, spelling, word choices and correctness (ibid.). This is also evident in the discussion forum analysed for this study as there plenty of incorrect forms and words spelled in different in ways in different posts.

In addition, discussion forums most often have a specific purpose for which they were meant and a specific target audience. While the linguistic data from discussion forums is easy to access, save and archive for research, the researcher needs to make methodological decisions at an early stage in regard to the degree of contextualisation and familiarisation (Kytölä 2012, 110). Understanding the community aspects (beliefs, values, favoured topics, nuanced meanings, ways of writing) of the forum complement the analyses, descriptions and interpretations of text, even if the researcher only focuses on multilingualism aspects (Kytölä 2012, 110). In addition, an ethnographic understanding of the forum can aid the researcher in making ethical decisions regarding the usability and anonymisation of the data (ibid.). For example, the usernames of the forum users should be omitted from the posts that are being analysed for the study. In addition, any personal data about the user if there is any for one reason or another depending on the forum should not be published. In this study, usernames are anonymised by pseudonymisation.

3 Previous research on codeswitching in electronic writing

In this section, I will discuss several studies that deal with codeswitching in written contexts. There are only a few studies here, but for the purpose of this study it is enough since there are only a small number of research papers that focus on analysing written codeswitching, especially in discussion forums and the works discussed here provide enough information for conducting my own thesis and analysing results.

First, a study by Kortelainen and Kolehmainen (2022) focuses on the forms and functions of multilingual practices in an online discussion forum that is meant for German-speaking migrants in Finland. It explores Finnish codeswitches and the alternating use of German and Finnish in German forum messages. Their research questions are: “[...] how Finnish elements are structurally embedded into an otherwise German syntactic frame in this online forum and on the other hand, what communicative functions the embedded Finnish codeswitches serve in the forum discussions” (Kortelainen and Kolehmainen 2022, 8). The data consists of 179 discussion threads in the *Deutsche in Finnland* (‘Germans in Finland’) discussion forum. The age of the posts varied, with the oldest being from 2006 and most of the others from 2010-2018. The posts were stored as word-files and then read through and manually searched for Finnish codeswitches, which were copied and saved into an Excel file.

In total, 616 codeswitches into Finnish were found and they were analysed on the basis of previous research and taking into account their structural and functional properties. Only 6% of the analysed switches consisted of intersentential codeswitches which involved complete Finnish sentences, and they mostly consisted of quotations that were directly copied from other web sites or excerpts of reported speech (Kortelainen and Kolehmainen 2022, 19). 82% of the switches consisted of intrasentential codeswitches and 12% consisted of metalinguistic commentaries, which were treated as a distinct category since they are “morphosyntactically independent and do not thus provide information on the structural properties of German-Finnish codeswitching” (Kortelainen and Kolehmainen 2022, 20). 63% of intrasentential switches were single common nouns, and these were often culture-specific items or concepts that were specific to Finnish culture, and they often had German articles or pronouns (ibid.). In some instances, common nouns were capitalised as both common and proper nouns are capitalised regardless of their position in German, and in some cases, they were written with a lower-case letter, which followed the Finnish norm (Kortelainen and Kolehmainen 2022, 21). In addition, Finnish nouns received the gender of the synonymous German word (for

example, *letku* ('hose') appeared in the masculine gender and *putki* ('tube') in neuter gender due to German counterparts being masculine and neuter) (Kortelainen and Kolehmainen 2022, 21–22). There is variation in some instances as for example, the gender of the word *viili* varied between masculine, feminine and neuter in the study, which might be due to the lack of a German synonymous word for the Finnish lexical item, which in turn causes indecisiveness about the correct gender (Kortelainen and Kolehmainen 2022, 22)

Moreover, plural marking was also involved in the integration of Finnish nouns in German sentences. Kortelainen and Kolehmainen (2022, 22) point out that only 16 Finnish nouns were in plural. The noun received the German plural ending in three cases, in one case the noun had no plural ending and the rest received the Finnish plural marking (*ibid.*).

Kortelainen and Kolehmainen (2022, 23) conclude that the use of articles, gender assignment and plural marking of nouns belong to features that can show the differences between spontaneous oral data and written data, because in written texts the writer may reflect and edit their linguistic choices before posting their message, which is not possible in oral communication, although Kortelainen and Kolehmainen note that more data is needed for a more detailed analysis.

Proper nouns were the second largest category of intrasentential switches (17.2%) and these included mostly names of referents typical to Finnish society in which the forum members live and exonyms (toponyms, like *Saksa* and “special geographical names used by people outside the place”) (Kortelainen and Kolehmainen 2022, 23). Exonyms have German equivalents while other proper nouns do not (*Helsingin Sanomat*) have. Kortelainen and Kolehmainen (2022, 23) state that “[t]he use of a Finnish exonym in a German sentence instead of the German toponym Deutschland ‘Germany’ is a signal that Finnish is – alongside the matrix language German – considered a common resource, the other we-code of the group”. Other intrasentential switches that were less common belonged to the following categories: noun phrases, interjections, bilingual German-Finnish compounds, single adjectives, adjective phrases, verbs and “other” (Kortelainen and Kolehmainen 2022, 20).

The results of the study show that “Finnish language” (37.2%) and “culinary corner” (25.2%) were the discussion categories that had the largest amount of codeswitches while “arts and culture” (0.3%) and “children and family” (2%) had the least (Kortelainen and Kolehmainen 2022, 26). Central functions of the Finnish codeswitches included: cultural switches (concepts

and objects specific to Finnish culture), quotations (codeswitching as a contextualisation cue marking “voices” in writing), greetings (also farewells and other well-wishes), reiteration (modified repetition of a message first expressed in German), linguistic errors, slips of the tongue and puns and metalinguistic commentary (discussion of Finnish grammar) (Kortelainen and Kolehmainen 2022, 25). They also noted that cultural switches are often not explained in the forum, which could be viewed as Finnish being used as a we-code for the participants in the forum since knowledge of the Finnish language and culture is required (Kortelainen and Kolehmainen 2022, 29).

Kortelainen and Kolehmainen (2022, 37) note in their conclusion that their analysis shows that the functions of codeswitching identified in previous research also apply to the switches they studied, and they also note that the use of Finnish cultural items was sometimes motivated by a lack of German equivalent, but they also functioned as contextualisation cues. In addition, there is a dependency between the topic that is discussed, and the frequency of Finnish codeswitches (*ibid.*). They also argue that “[i]n the context of the investigated Finland-themed forum, the use of Finnish as a common communicative resource can be interpreted to function as a tool of creating a sense of an in-group identity” (Kortelainen and Kolehmainen 2022, 38). Forum members seek common ground and signal their membership by using Finnish codeswitches that “carry the expectation of shared multicultural and multilingual experience”, but the study does not intend to find out what the codeswitching practices reveal about the real-life identities of forum members (*ibid.*). Whether the written codeswitching practices are similar to multilingual practices in the spoken context and to what degree they can be generalised to represent offline multilingual practices remains unknown (*ibid.*).

Another study that focuses on codeswitching in written form has been done by Kristin Vold Lexander (2012) and her research focuses on multilingualism in text messages of students in Senegal. The participants consisted of seven women and eight men who were of similar age and socio-economic background, but all of them belonged to different ethnic groups and all had different language skills, although all of them spoke Wolof and French (Lexander 2012, 149–150). Communication in the group and with others outside the group was studied. The data consisted of 450 text messages, and it was collected during three period of fieldwork in Dakar, from 2005 to 2007. The study focused both on linguistic features and social aspects and how multiple languages can be used to signal different identities or different aspects of identity in writing (Lexander 2012).

The comparison of computer-mediated communication with other written or spoken language practices makes it possible to study the written use of codes that are typically reserved for spoken interactions (Lexander 2012, 149). According to Lexander (2012, 158), SMS messages are not only read, but also seen. The corpus of the study contained very few emoticons, but visual effects like capitals (can also be used to flag codeswitching) and quotations are used in combination with codeswitching (ibid.).

Computer-mediated texts and spoken language share some of the same features, and informal computer-mediated communication (like discussion forums) utilises the same registers as informal spoken conversations, and instant messaging is similar to conversation with “short answers, quick exchange and references to body language” (Lexander 2012, 149). Language choice and codeswitching are also visible in written language and can be combined with other visual features of texts (ibid.). Moreover, the analysis of codeswitching is enhanced if we take into consideration the interaction between multiple visual characteristics in informal computer-mediated communication and language mixing (ibid.). One of these visual features is orthography, where abbreviations and respellings are often the most typical aspects being studied in computer mediated communication (ibid.). “Orthography can be used to signal language attitudes, contrast languages or to bring them closer together” (ibid.). Lexander (2012, 161) also points out that “[t]he ‘academic’ style and codeswitching is expressed through another visual feature, the orthography of Wolof: the writer uses *u* (instead of the more often used French-like orthography *ou*) and *x* (instead of *kh*), signalling his familiarity with the official transcription of Wolof”.

The switching of codes also takes place within languages and not just from one language to another as informants switch between different registers and Lexander points out the difference between *français correct* and *français déformé* (‘deformed French’) in her study (ibid.). The orthography between the two registers differs as *Français correct* is close to standard French in writing while deformed French contains abbreviations that are non-formalised, numbers instead of letters and respellings (ibid.). Young people use *français* correctly when messaging people older than the writer and when messaging authorities while deformed French is used between young people and using more standard language is considered strange (Lexander 2012, 161–162).

In the conclusion Lexander (2012, 165) argues that visual features (For example, capitals, quotation marks) together with codeswitching are used for various purposes, like classifying

codeswitching, but notes that codeswitching appears as a visual feature itself. Lexander (2012, 165) also states that “[t]he use of ethnographic tools for data collection reveals motivations for language choices, related to managing interpersonal relationships or to identity aspects”. In addition, orthography is an important part of each text as it is used for defining codes as well as switching between codes, and the comparison of codes in language practices helps us to understand the dynamics of sociolinguistic situation better (Lexander 2012, 166). Lexander (ibid.) also notes that studying the relationship between smileys and codeswitching or a comparison of the functions of emoticons and switches in monolingual and multilingual corpora would be an interesting object of research since these aspects could not be studied in the study.

A study by Nur Syazwani Halim and Marlyna Maros (2014) focuses on the occurrences of codeswitching in electronic writing. The study investigates the functions of codeswitching by analysing Facebook status updates of five adult Malay-English bilingual users who are English educators teaching in tertiary institutions in Malaysia (Halim and Maros 2014, 129). The data included 439 Facebook status updates that contained codeswitching, and the data was analysed before being quantified according to frequency and classified based on the functions (ibid.). The functions of codeswitching discussed in the study are: switching for quotation, addressee specification, reiteration, message qualification, clarification, emphasis, checking, indicating emotions, availability, principle of economy, simpler English or Malay words and expressions and free switching (Halim and Maros 2014, 129–133). Halim and Maros (2014, 133) conclude that codeswitching is used for demonstrating competence in both English and Malay. In addition, it is used as a strategy to communicate and negotiate meanings and to achieve communicative and stylistic objectives and enhance interactions (ibid.). Furthermore, codeswitching “should be perceived as a functional sociolinguistic or communicative tool that adds colours to both spoken and written communication in multiracial community rather than a lingo that leads to the corruption of the pure language” (ibid.).

The previous studies discussed here were helpful for choosing a method for my study and how to analyse and classify codeswitches in my materials.

4 Research methodology

The purpose of this chapter is to first introduce the discussion forum *Complete France Forum* in 4.1 from which my materials were collected. This section also introduces how the materials were collected. In 4.2 I explain how the materials were classified and analysed.

4.1 Materials

Complete France Forum is an online discussion forum that allows its users to ask for advice or help on different matters, such as legal and finance issues, moving or travelling to France or house renovations. The forum does not offer information how many people have made an account on the forum, but most likely there are at least a few thousands registered users if not more as the high number of discussion threads and posts suggests that the forum is an important tool for a lot of people and new posts and discussion threads are created often.

Users of the forum can discuss and share their experiences on many different topics, such as French cuisine, language, literature, living and education. Users can also discuss many other things that do not have anything to do with France, but mostly only in the general discussion category. English acts as the main language of the forum as the majority of the discussion threads are in English, but many posts can contain some other language (most often French) or multiple languages in one post. There is an assumption in the forum that most of the users also understand French or use a translator if they do not understand the language since when French phrases or acronyms are used, they are never translated or explained in English.

Usually, the parts that are in French in a post are still relevant for understanding the meaning of the post. One can view all the topic threads and posts without creating an account in the forum, but an account is required if one wishes to create new topics or write a comment in an existing topic thread. The forum has moderators who supervise the forum and remove any posts or topic threads that are against the rules of the forum. For example, posts that are harmful, discriminatory (against race, religion, gender etc), threatening or offensive, are deleted. These rules are part of the Terms of use guide and the user has to agree to those rules if they wish to create an account. Normal users are not allowed to create new topic categories or sub-categories as probably only admins of the forum have that right.

The forum went through some changes in 2021 as it has a new software. Users who had never created a post in the forum had to create a new account as only users who had made a post in the past were migrated in order to get rid of bots and automatically generated accounts. The

URL of the forum also changed but visiting the old URL redirects to the new. All the forum posts were migrated from the old forums. This did not have an impact on my study as the materials were collected after these changes.

The forum is part of a larger site called *CompleteFrance*. The site has a wide variety of content, such as travel guides, property listings and information on the different regions of France. It is owned by France Media Group business, which was established in 2003 and they bought the site that the forum belongs to and several other magazines from Archant in 2022. France Media Group is a business with products and services serving the international Francophile community and their headquarters are located in Bath, UK and they also have satellite offices in France. The business has grown to become the world's leading publisher of print media, digital media and events for Francophiles.

The forum is divided into 11 topic categories¹:

- 1) About this forum
- 2) Moving
- 3) Legal and Finance
- 4) Building and Renovation
- 5) Living
- 6) Owning/Running a business
- 7) Leisure and Pastimes
- 8) French Culture
- 9) Travelling to/from France
- 10) Holidays
- 11) General Discussion
- 12) Regions

The first discussion category is excluded from this study since it has only two posts created by the administrators of the forum and provide general instructions about using the forum and other users are not allowed to create discussion threads or reply in this category. Most of the discussion categories also have at least several subcategories. The *Leisure and pastimes*

¹ The discussion forum and its various discussion categories can be found at:
<https://forum.completefrance.com/index.php>

discussion category has the greatest number of subcategories (11), followed by *Living* (10) and *Regions* (8). The *Regions* category has its own subcategory for discussing topics specifically related to each region of France (for example, *Paris Ile de France*), *West* (*Bretagne, Pays de la Loire*) and others. Most of the subcategories there have only around 4000-6000 posts (*South West* has the most while *East* has the least (only 276 posts)). *Holidays* and *Building and Renovation* (around 49 000 posts) both have only one subcategory and *Holidays* category also has the least number of posts (around 1900). The forum does not list the exact number of if it is over one thousand, but rather an approximate number of posts that each discussion category and subcategory have by rounding it up or down (for example, 4520 would be 4,5k). In addition, the forum shows how many times each discussion thread has been viewed, which indicates how many users have been interested in the topic so far. The popularity of each discussion category varies as the number of posts in each discussion category varies greatly. *General Discussion* has the greatest number of posts (around 212 000) and *The Complete France Post Bag* (94 000) and *Other Topics* (81 000) are the most popular subcategories. The first subcategory is generally about everything about France and all things French, where the discussion topics can be more varied. *Other Topics* is specifically for topics that do not relate to France and almost anything can be discussed. There is also a *Forum Technical Support* subcategory that is meant for any technical issues the users might encounter users can propose new ideas regarding the forum.

Living (around 164 000 posts) discussion category is also quite popular. It includes subcategories, such as *Driving in France*, *Health*, and *Pets*. *French Satellite TV*, *French Internet and Telephone* (47 000) is the most popular subcategory. Interestingly, the forum claims that *French Education* subcategory has no posts at all, but if you open the subcategory then you can see that there are a lot of discussion threads (around 700) so that might be due to a technical error. The same problem applies to a few other sub-categories in the forum (for example, *Books and Literature* in *Leisure and Pastimes*). It should be noted that topic categories still show the correct total number of posts in that category and the number information is only missing when viewing the sub-category. *Travelling to/from France* has 20700 posts, but when you open the topic category, you can only see that one category has 19400 posts while the other subcategory claims that there are no posts there, but by opening it one can count the remaining posts.

In general, based on the number of posts across the topic categories and how many times different discussion threads have been viewed it can be said that the discussion forum is quite

active, and it is an important platform for many people to discuss different aspects about France and deal with France related problems by asking others. Several topic categories are much more active and have more discussion threads than some of the other topic categories where there are not many discussion threads and many discussion threads are quite short, and in some subcategories, no one has created a new discussion thread for several years. Based on the analysed posts, some users post much more frequently than others and are very active across different discussion threads while other users participate in the discussion only a few times. For instance, 19 of the 150 analysed posts that contained codeswitching were made by one user. This may suggest that this particular user likes to codeswitch more often than other users. It is possible that the user uses French more in his/her everyday life or prefers French words. The users of the forum come from different countries around the globe and have different language skills and most of them use English as a lingua franca. For that reason, it must be noted that the data of this study does not come solely from L1 speakers of English with French as their L2, but this does not pose a problem for this study as each user's L1 and L2 knowledge is not relevant for this study. It is only important that the users converse mostly in English and switch into French occasionally. Unfortunately, it is impossible to determine the nationality of users and their first language unless that information is specifically provided by the user himself/herself in certain contexts. Nevertheless, it can be easily observed that most of the users can write in English, although the proficiency levels of the users can vary on the basis of how they write. A great number of users have at least some knowledge of French, which can be observed from posts that contain a lot of French codeswitching, or their familiarity with the language is visible in some other way depending on the context. The contents of some discussion threads suggest that at least a handful of users may be from the UK. There are quite a few posts in various threads that are completely in French. In these cases, users sometimes reply in English to these French posts and sometimes in French, but in both cases, it is clear that the users who reply to a French post understand French well. There is also a discussion thread named "Qui veut "chatter" en francais?" ('Who wants to "chat" in French? ') that has around 1800 replies where people chat mainly in French, although there are some brief phrases in English sometimes. The length of the thread and the number of users posting solely in that topic provides evidence that there are plenty of users who have a good knowledge of French. It is likely that there are at least some native French speakers on the basis of the information that the users provide in various discussion threads and how they write. However, it is not possible to give an exact number since the information is not verifiable.

The topics of the discussion forum can be categorised into the following categories in terms of languages:

1. The posts are completely in English with no codeswitches to another language. This is the most common category since the majority of the posts are only in English as it is assumed that everyone on the forum understands English. This assumption is also proven by the fact that the rules of the forum, instructions on using the forum and terms of use are only available in English. The users of the forum are from different countries with different first languages and language skills, but English is the lingua franca on the forum which enables communication between different users. In addition, this makes the forum more accessible to people around the world and more people can participate in the discussion which is the point of a discussion forum.
2. The posts contain both English and French. Many posts were mainly written in English with occasional French codeswitching and some posts were mainly written in French with occasional English codeswitching. I consider a post to be mostly in French if over 50% of the words in the post are in French. This is the second most common category, and the number of posts mainly in English with minor/occasional French codeswitching is much larger than the number of posts mainly in French.
3. The posts are completely in French. French-only topic threads are rarer than the previous categories, but they do appear in a few discussion threads (for example, one discussion thread is specifically meant for using only French). Some posts in various discussion threads might be in French, even though all other posts in the thread are in English which is also a type of codeswitching, but not considered in this study. The topic category *French language* was not included in the study because it contained posts only in French. The
4. The posts are in some other language than English or French, like Italian, although this category is very rare. I found no discussion threads that were fully in some other language or even a post fully written in other language. Mostly, these are different kinds of short phrases (for example, greetings) in some posts around the forum.

For this study, only posts that are primarily in English and contain French codeswitching are considered since posts entirely in one language only do not contain any codeswitching and other languages than English or French are not relevant for this study. Especially posts that belong to category three or four mostly belonged to discussion thread that were excluded from

the study while gathering materials. In addition to forum posts, titles of the discussion threads are also taken into account if they contain codeswitching. If the title is a short phrase (2-3 words) and in French completely, then I still consider it to be a codeswitch if the topic creator's first post is mostly in English (for example, *le repas anglais*)

4.2 Methods

The research focuses on the public part of the forum where all discussion threads and posts are visible for everyone, even for people who do not have an account in the forum. Due to that, no research permissions were required from the admins or moderators or from the users whose posts were analysed in order to use the posts as research material. I did not create my own account in the forum, so I have not influenced the discussion or codeswitching in any of the discussion threads in any way. In examples that are used in this study, the users are anonymised by pseudonymisation. In addition, the examples cited in the study are neutral in tone so that they do not cause offense to anyone. The focus of the research is on the posts of the forum and discussion thread titles only, so usernames of the forum users are not taken into account. Furthermore, quoted posts (a user has quoted the text of another user) are excluded from this study as that prevents analysing the same post and its codeswitches multiple times.

The data was collected in the time period October 2022- August 2023. Posts that contained codeswitching were collected from 12 discussion threads in total. The discussion categories were from the following topic categories: 1) *Moving*, 2) *Legal and Finance*, 3) *Living*, 4) *Leisure and Pastimes*, (2 discussion threads) 5) *French Culture*, (2 discussion threads) 6) *Holidays*, 7) *General Discussion*, (3 discussion threads) 8) *Regions*. More than one discussion thread was picked from a few topic categories because these categories had a lot of different discussion categories and the themes discussed in them were quite varied, particularly in the *General Discussion* category. In addition, they contained plenty of codeswitching in different contexts, which is essential for my study. No discussion threads were taken from *Building and Renovation*, *Owning/Running a Business* and *Travelling to/from France* categories. Many discussion threads in those categories were quite short (only a few posts) and many threads did not contain any codeswitching or one or two codeswitches were used repeatedly or only a very few switches were found in a discussion thread. In addition, many discussion categories that had more than a few replies were very old. It is likely that some discussion categories in those categories contained a significant amount of codeswitching, but analysing every discussion thread in the forum would not be feasible for this study. In total, 582 posts were

thoroughly analysed for codeswitches across 12 discussion threads and 150 posts in total contained codeswitching and were copied and saved into a Word file. Many other discussion threads were also analysed, but they were excluded from the study for various reasons. For instance, discussion threads that had only a few posts (3-4) or no replies at all were not analysed as they generally did not contain any codeswitching or very little. Discussion threads that had over 5 pages of posts (over 100 posts) were also excluded as their analysis in order to capture more varied range of codeswitches as in many long discussion threads only one or two particular codeswitches occurred many times in the same form. This would have resulted in too many codeswitches and posts to be analysed for this study. The oldest posts in the analysed discussion thread are from the year 2005 and the newest posts are from the year 2022. Most of the analysed discussion threads range from 2014 to 2020.

The analysed discussion threads were read through and manually searched for posts that contained French codeswitching. Every post that contained codeswitching was copied and saved into a word file. In many cases, I copied and pasted the entire post from the forum that contained codeswitching. However, in some cases, I removed paragraphs or longer sentences that did not contain any codeswitching if it was not relevant for the context in order to make it easier to analyse each post on its own and not having too much surrounding text. Every French codeswitch in any post was bolded and analysed together with the surrounding text and previous posts if necessary. The structure and functions of codeswitches were analysed while also taking previous research into account. The structure of the codeswitches was analysed first and the switches were divided into intrasentential, intersentential or metalinguistic commentary categories. The same categorisation was also used in Kortelainen and Kolehmainen's study (2022, 19), which was a good basis for my study. After that, the switches were categorised into nouns, adjectives, verbs or something else similar to Kortelainen and Kolehmainen's study, although with a few modifications (for example, formulaic expressions and prepositional phrases) which better fit my study. Finally, the codeswitches were analysed in terms of their function in a context and divided into categories, like culture, quotations and humour. These categories were inspired by previous research that analysed functions of codeswitching, like Gumperz (1982), Kortelainen and Kolehmainen (2022) and Halim and Maros (2014). In addition, language discussion category is added by me to this study since many codeswitches fit better into those categories in my thesis than the categories used in previous works.

5 Analysis

5.1 Structural analysis

The structural analysis of French codeswitches provides information on how English and French are mixed in the French themed discussion forum posts. Only intrasentential codeswitches are discussed in more detail in the structural analysis since intersentential consisted of quotations that will be discussed in the functional analysis. Table 1 shows how frequently different types of codeswitching appear in the discussion forum. The focus is on the token number of codeswitches, not type numbers, meaning that each individual codeswitch in every post is counted, even if it occurs multiple times.

Table 1 Structural properties of French codeswitching

	n	%
Intrasentential codeswitches:	249	82
Intersentential codeswitches:	25	8
Metalinguistic commentaries:	31	10
Total:	305	100

Table 1 shows that codeswitching between English and French in the *CompleteFrance* discussion forum consists mostly of intrasentential codeswitches (82%) as 249 switches out of 305 belong to that category. French words (for example, *salle* and many others) and phrases varying in length (from changing just one letter (*cinéma*) or short noun phrases, such as *les filles* to longer phrases, like many French sayings and idioms (*c'est la vie*) or organisation names) belong to intrasentential codeswitches. Only 25 switches out of 307 (8%) belong to intersentential codeswitching, which mostly consist of quotations from other sites (for example, quotation from the president of France, song lyrics, and statistics from other sites) and these varied greatly in length. Some were short ones by just having one short sentence completely in French while other quotations had several paragraphs only in French (for example, copy pasting an entire article). Several instances of reported speech in the data also belong to intersentential codeswitching. Kortelainen and Kolehmainen (2022, 31) state that “[r]eported speech encompasses quotations of speech and writing produced either by the writer themselves or another person at an earlier point in time”, which I also used as the criterion for reported speech in my data. The following example contains reported speech:

The remaining 31 switches (10%) belong to the metalinguistic commentary category. In Kortelainen and Kolehmainen's study they are treated as a distinct category "[...]as they are morphosyntactically independent and do not thus provide information on the structural properties of German-Finnish codeswitching" (Kortelainen and Kolehmainen 2022, 19–20). In many cases, they consisted of different phrases or sentences that had multiple parts of speech and so were not part of the structural analysis.

There were only a few single nouns, noun phrases and adjectives that were taken into account in the structural analysis. Metalinguistic commentary category mostly consists of switches which involve discussion about the French language itself and comparing the different meanings of word certain words between English and French and all of them belong to the *language discussion* category which will be discussed in the functional analysis section.

Table 2 demonstrates which parts of speech and types of phrases appeared the most often in my data and which ones the least. These consisted of intrasentential switches since intersentential switches and metalinguistic commentary categories often had phrases or sentences with multiple parts of speeches. Each category is discussed separately, and examples are provided.

Table 2 Parts of speech and phrases in intrasentential codeswitching

	n	%
Single nouns	131	54.1
Proper nouns	43	17.7
Noun phrases	31	12.8
Adjectives	11	4.5
Adverbs	5	2.0
Interjections	4	1.7
Prepositional phrases	4	1.7
Verbs	2	0.8
Formulaic expressions	11	4.5
Total	242	100

5.1.1 Nouns

Table 2 shows that nouns were the most frequent codeswitches in the data. 131 out of 242 codeswitches were single common nouns (54.1%). Single nouns referred quite often to culture

specific concepts and things or terms and words that are mentioned many times and common in French society. These include for example, different foods (*Parfait* ‘frozen dessert’), places (*boulangerie* ‘bakery’), French society concepts (*Notaire* ‘legal officer’) and professions (*conciergerie* ‘doorkeeper’). Single nouns also could refer to common everyday life objects and things (*tapis roulant* ‘moving walkway’, *crédits* ‘credits’ etc). Many of these single common nouns are not specific to French culture and society. Generally, if nouns did not belong to the culture category, then they belonged to “free switching” category, meaning that the writer had his/her own reasons for choosing the French words instead of the English one. This is especially interesting in words, such as *crédits*, *déductions* and *région*, because the only difference between the English and French word is just one letter (English *e* vs French *é/è*). It is possible that French is the second or first language of the writer or he/she just prefers the French form. In addition, French single nouns are often accompanied by English articles (the *facture* ‘bill’) or pronouns (my *ambulancière* ‘paramedic’) or something else. precedes the noun. French articles and pronouns also appear sometimes, but not as often as English ones. It should be noted that while English articles were used before the noun and English has no grammatical gender, the French noun itself followed French grammar by appearing the correct gender (masculine or feminine), like in the following example:

- (1) My sister (again) is a professional **liberale** and probably earns the average salary but she is massively hit on tax because of her business regime (User 1).

In example 1, everything preceding the noun (*liberale* ‘profession’) is in English, including the indefinite article “a” and adjective, but the noun appears in its feminine form (masculine form would be *libéral*) as it refers to the sister (which is a feminine noun in French) despite not having a feminine indefinite article (*une*). Kortelainen and Kolehmainen (2022, 20) in their study note that Finnish single and proper nouns in German sentences can be accompanied by German articles and pronouns which express grammatical relation and due to that, there is no need to inflect Finnish nouns in German sentences. However, English and French are a very different language pair when compared to German and Finnish in terms of grammar and other areas. Finnish nouns have no gender or articles while German nouns have both (masculine, neuter, feminine). In my study, both English and French use articles and French nouns always have gender (masculine or feminine), and French nouns are inflected. Kortelainen and Kolehmainen (2022, 21) state that in their study Finnish nouns receive the gender of the equivalent German noun while in my study, French nouns often receive the

English indefinite or definite article, but the noun itself appears in the gender that it has in French or if it can appear in either form then it depends on the premodifier.

The plural marker is mostly the same between English and French as both use -s the most, but French also has two plural articles (*des/les*) (*les élections* ‘the elections’). It should be noted that there are exceptions in French regarding plural marking. For example, some nouns form their plural with an -x (*un journal- des journaux*). French nouns that end in -eau, -au or -ou receive the -x plural marking. In addition, nouns that end in -s, -x or -z in the singular do not change in the plural. (*un pays-des pays*) and there are also irregular plural forms. However, none of the nouns in my data had a -x plural or irregular forms as all of them had the -s form. It is possible that the other forms appear in other discussion threads that were not analysed for this study as they are not rare in French.

Capitalisation patterns in the data also had some interesting results as there was variation between French codeswitches. In general, many words that are capitalised in English, but are not in French. For example, weekdays, geographical terms, language and nationalities are not capitalised French. Here are some examples from the data: local *Mairie*, *Notaire*, *Gites*, *gites*, *anglais*. First, *Mairie* (In larger towns and cities, it is known as the *Hôtel de Ville*) refers to a town/city hall, but it can also refer to a city council, mayor’s office or the city as a political entity. The noun appears four times in the data and each time it is capitalised, although it should not be capitalised in French. In other internet sites, it generally appears uncapitalised. It is possible that it is capitalised as the users apply an English principle to it or to highlight as a word from another language. Similarly, the noun *notaire* appears twice in the data and it is capitalised, but normally it is not. *Notaire* is a French public officer, who assists people in many situations, such as buying a house or getting married. The noun *gite* appears 10 times in the data, and in two cases it is capitalised. It refers to a type of accommodation in France. This might be just the writer’s preference to capitalise foreign words as both times it is capitalised by the same user and the same pattern could also be observed with other French codeswitches in posts made by the same user.

The codeswitch *anglais* (for example, *repas anglaise* ‘English meal’ or “those *anglaise*” ‘those English people’) appears several times in the data and it is one clear example of the codeswitch following French grammar rules as it is always uncapitalised in the data as nationalities are not capitalised in French. In addition, there was variation in capitalising different foods. Some foods should have been capitalised but were not (for example, *crepe*

suzette as *suzette* is a proper noun). Similarly, the yellow vests movement (*gilets jaunes*) was mentioned plenty of times in one discussion thread. It appeared both in capitalised and uncapitalised forms appeared several times.

In general, the capitalisation of French codeswitches depends on user's preferences in the forum as many nouns appeared in both forms and did not seem to follow the norm of either language. However, some switches did follow French grammar rules for capitalisation. Interestingly, there were no capitalised codeswitches which would be capitalised in English, but not in French. One reason for variation and use of incorrect forms might be that people on the forum write fast and do not pay that much attention to grammar, which is common of social-media type of writing. The forum is not a very formal setting, and nobody points out minor language mistakes of other people so using correct grammar is not a high priority in the forum if everyone can understand what is said. The variation might also be due to different language proficiency and education levels among the forum users.

Gender, capitalisation, articles and plural marking features in the written data can be different from spontaneous oral data as in the written context, the writer has the opportunity to edit their linguistic choices before posting (Kortelainen and Kolehmainen 2022, 23). In the discussion forum, the user can decide how much time goes to writing a post, when to codeswitch and proofread it many times if necessary (and edit it even after posting in a discussion forum). The reasons for appearance of incorrect forms in codeswitches can also be different between oral and written data as their appearance in written context suggests that the writer does not care about grammar that much, the post was written quickly without editing, the writer does not know the correct form or there are personal reasons.

5.1.2 Noun phrases

There were plenty of noun phrases as 31 out of 245 (12.8%) codeswitches belonged to this category. Noun phrases mostly consisted of French nouns that were preceded by an article, adjective or something else in French. These include for example, *repas anglais* ('English meal'), *les filles* ('the girls'), *un étrangère* ('an outsider') and *petite rue* ('small street'). In French, adjectives can sometimes appear after the noun, which can be observed in the following codeswitch: *Centres Commerciales* ('shopping mall/center'). In general, noun phrases also refer to general and everyday words and concepts. Many noun phrases belong to the food or society categories, and a few refer to places. There are very few nouns in the data that have a French article. This suggests that people prefer using English articles before a

French codeswitch noun, which might be due to the theme of the forum (France) or personal preference. Another observation worth noting is that in several noun phrases when referring to English people in general or some food, *anglais* is preceded by a French word or article (*repas anglaise, cuisine anglaise, les anglaise*) instead of “the English”. Similarly, the word “outsider” does not appear at all, but instead the French equivalent appears three times (*un étrangere*). This could be due to the attitudes between English and French and using French codeswitches may serve as a way to distance oneself from the other culture. In addition, French people seem to often refer to English as “anglais” so it might also be a habit for many people and seen as normal.

5.1.3 Proper nouns

Proper nouns were the second most frequent category after single nouns as 43 out of 245 switches (17.7%) were proper nouns. Proper nouns differ from common nouns, and they often consist of expressions that are transferred without any changes from one language to another (Raukko 2017, 105–106). In the data, proper nouns mostly referred to specific places (*Côte d’Azur, Parc de Ecrins*), organisations (*Médecins sans frontières* ‘Doctors Without Borders’, *INSEE* ‘national statistics bureau in France’) and all kinds of different French social and society systems that one living in France has to deal with at least at some point. These include for example, *SDF* (Personne sans ressources ‘homeless person’), *SMIC* (‘minimum wage’) and *CV* (carte vitale ‘health insurance card’). Some switches also refer to French newspapers (*La Figaro*), French internet sites (*Le Bon Coin*), French companies (*Direct Énergie*) and French political parties (*Le Front National*). Most of these switches do not have English equivalents, but some do, like *Médecins sans frontières*, which would suggest that French proper nouns are preferred by at least some forum users over the English ones. In addition, the different acronyms referring to different social systems and institutions are almost never explained (for example, *CV* refers to health insurance card or *CAF* to state welfare agency) meaning that there is an assumption on the forum that everyone is already familiar with these different acronyms and French systems.

5.1.4 Other categories

Other structural categories occur less frequently than nouns. This can also be observed in Kortelainen and Kolehmainen’s study (2022). There were only 11 codeswitches (4.5%) that belonged to the adjectives category. Only a few adjectives belong to metalinguistic

commentary and language discussion categories since they involved discussion about using an adjective correctly in a certain context. Some belonged to society category (for example, *Régional* elections), and a few belonged to *free switching* category because the French adjectives were probably in French due to writers' own preferences, which can be observed in the following example:

- (2) Autoroute travel in our area has become '**gratuit**' as the exit barriers at the **peage** are permanently raised and a representative of the '**gilet jaune**' collects the ticket, presumably to pass on to Macron to enable the companies who operate the autoroutes to reclaim the fee (User 2).

The adjective *gratuit* ('free') does appear in the *Oxford English Dictionary*, but it is labelled as rare and obsolete (OED, s.v. "gratuit," adj). Furthermore, it was only used in 16th and 17th century in the English language. However, the adjective is common in French, so the adjective is considered a codeswitch for this study. Here the French adjective was probably picked due to the French context and his/her own preference. The writer has also used single quotation marks for the codeswitch in order to highlight as a word from a different language. In general, the presence of French adjectives was very small, and people seem to prefer using English adjectives. In addition, there are two other codeswitches in the example (*peage* 'toll' and *gilet jaune* 'yellow vest movement').

The presence of adverbs was also minimal as there were only five adverbs in the data (2%). These included for instance *enfin* ('finally', 'at last'), *amicalement* ('with kind regards', used as a farewell in letters/emails) and *à deux* ('being between two persons in intimate relationship'). In addition, the presence of prepositional phrases was also minimal as there were only four such phrases. The prepositions that appeared in the data were: *chez* ('at' or 'in'), *sans* ('without') and *à* ('to', 'at', 'in'). They were always followed by a French word (*sans sirop*, *chez toi*, *à doux*) and usually they occurred at the end of the English sentence. Only in one instance the codeswitch was separated from the sentence by parentheses (*sans sirop*).

The frequency of interjections was very rare as only four (1.7%) appeared in the data. Interjection is "[a] natural ejaculation expressive of some feeling or emotion, used or viewed as a Part of Speech" (OED, s.v. "Interjection," Grammar) These include: *beurk*, which is similar to 'yuck' and *miam miam*, which means 'yum-yum'. Both relate to food. In addition, *mais oui* is an interjection as it is a popular expression, and it means 'of course' (literal meaning: 'but yes').

Codeswitching involving verbs was the least frequent as only two instances of a single French verb were found in the data. It might be that switching to a French verb is more difficult as verb tenses and moods are different between the two languages and French verbs can vary drastically in written form.

- (3) Plenty of people do make a living but whether it is a good living **dépends** on their expectations I guess (User 3).

In example 3, the first verb was *dépends* which in this case is easy to use since it is almost the same form in English except for the accented -é. However, grammatically the example is incorrect since the correct form in the example would be *dépend* (without the -s). It could be considered a hybrid form containing elements of both languages (English s-ending and French é). The usage of that form might be due to English interference or a simple unintentional language mistake. It might also be a conscious choice. The second verb was *detendez-vous* which means ‘relax’ or ‘chill out’:

- (4) You seemed to think you'd get insults..... all I can say is **detendez-vous!** Life is too short! (User 4).

In example 4, another French feature is that the verb is followed by a personal pronoun (*vous*).

Finally, I also included the category of formulaic expressions in the structural analysis. They are separate from the other parts of speeches, such as nouns and adjectives and consist of multiword units or phrases. They are not included in the functional analysis section since they do not fit into any of the categories there, but they are still an important part of language. There were eleven formulaic expressions in the data (4.5%). Quite a few of the expressions involved wishing good luck to someone as these occurred several times (*bon courage*, *bonne chance*) or telling someone to get well soon (*Bon rétablissement*). Other expressions included for example, *Bonjour la France* (‘Welcome to France’), *pas du tout* (‘not at all’) and *c’est la vie* (‘that’s life; oh well’). These expressions can be used in different contexts. Many of them are also well known internationally. Furthermore, they do not often have a specific cultural or humorous function.

5.2 Functional analysis

The purpose of this section is to analyse different functions of French codeswitches. Most of the French codeswitches correspond to the general functions of codeswitching which are

discussed by Gumperz (1982). The functions discussed in this study include: *cultural switches, greetings/farewells, quotations, humour, free switching and language discussion*. Many of these general functions also appear in Kortelainen and Kolehmainen's study (2022). *Free switching* function appears in Nur Syazwani Halim and Marlyna Maros' study (2014). A systematic quantitative analysis was not done for the functional analysis since the functions of codeswitching tend to overlap often, which is also noted by other researchers, such as Kortelainen and Kolehmainen (2022) and Yankova and Vassileva (2015). For that reason, providing an exact frequency of each function in the data is difficult and the results would not be reliable since many codeswitches in my data can also have multiple functions. Only approximate estimations are provided for most functions on the basis of rough tendencies observable in the data. Only the *language discussion* category has an exact number of occurrences since these codeswitches stood out clearly from the other categories.

5.2.1 Cultural switches

Most of the French codeswitches had a cultural function as over 150 cultural switches were identified in the data while all of the other function categories had under 50 switches. Especially most of the noun phrase and proper noun codeswitches had a cultural function. Cultural codeswitches refer to either different concepts and objects that are specific to French culture and society or in some other way associated with France. Kortelainen and Kolehmainen (2022, 27) state that “[a]lthough some Finnish items categorized as cultural, such as *lukio* (‘high school’), are not particular to Finland only, their Finnish referents nevertheless differ clearly from the referents of the corresponding German-language items”, which is also observable in my data:

- (5) Sounds like the cinema that I went to in Belgium to escape the rain and sleep off a hangover, first warning that all was not what it seemed was that the **concierge** was giving a hand towel to all the visitors that were 100% old men in trenchcoats who once in the **salle** were vigourosly polishing their glasses under their coats [:-)]) (User 3).

In example 5, the noun *concierge* exists both in English and French and it refers to a profession in both, but with some differences. It refers to a person who takes care of an apartment building or an usher in a cinema in French while in English it generally refers to a hotel employee. In the following example, *concierge* refers to a ‘usher’ in a cinema. In addition, there is another French cultural switch *salle*, which refers to a ‘movie hall’.

Kortelainen and Kolehmainen (2022, 27) note that “[t]hese culture-specific codeswitches can be viewed as cues that contextualize the referent as Finnish, convey its link to Finnish culture or society, and express its Finnishness”, and that also applies to French cultural switches in my data as can be seen in example 6:

- (6) We are permanent residents in France with **CV** and **mutuelle**. We had rented an apartment from a private owner for a holiday this week; it is a French registered business with a Swiss owner. The day before leaving I fell down the stairs and broke my leg. We have obviously cancelled the booking, but wondered if anyone has any idea about any avenue by which we might reclaim some of the cost of the apartment. We paid by virement from our French bank account (User 5).

First, ‘CV’ in this context does not refer to curriculum vitae, but to a ‘French health insurance card’ called *carte vitale*. Another cultural switch *mutuelle* has a literal meaning ‘mutual’, but in the example it refers to ‘health insurance’. In addition, the noun “virement” is a borrowing from French, according to the *Oxford English Dictionary*. By using these French cultural switches, the user on the forum specifically links the words to French health care system and not some other country. In addition, the frequent use of some cultural switches can be due to forum users being frequently exposed to them in their everyday lives (Kortelainen and Kolehmainen 2022, 27). For instance, many forum users who live in France likely have to often use terms, such as *Notaire* (‘public officer’) or *Mairie* (‘town hall’).

The cultural switches can also be divided into three different thematic categories (*cuisine, society and places*). The largest category was *society* as approximately 100 cultural switches belong to this thematic category. Switches related to *society* refer to different institutions, health care, professions, social security, organisations, companies, French internet sites and terms that relate to French society and everyday life in France. These switches include for example, *SDF* (*Sans domicile fixe*, ‘homeless person’), *gendarmerie* (‘military force’), *la poste* (‘postal service’), *chômage* (‘unemployment’), *Pole Emploi* (‘employment office’), *locataire* (‘tenant’) and *enfants naturels* (‘natural children’, a child whose father and mother were not married at the time of conception). The second largest category was *cuisine* with approximately thirty switches. These switches refer to different French foods, dishes, wine or to cooking in general. These switches include for example, *poulet fermier* (‘free-range chicken’), *terroir* (‘winegrowing term’), *sucré-salé* (‘salty sweet’), *le repas anglais* (‘English meal’). Finally, approximately thirty cultural switches were categorised as places, and these refer specific types of buildings, shops, parks and geographical places. These include for

instance, *gîte* ('furnished holiday house specifically in France'), *boulangerie* ('French bakery'), *leclerc* ('hypermarket chain'), *Parc de Ecrins* ('Écrins National Park') and *Salle de Fete* (community center').

5.2.2 Free switching

Approximately 35 codeswitches did not belong to any of the other function categories and were thus classified to free switching category. These consist of different French single words (for example, verbs, adjectives, adverbs, single nouns) or short phrases (for example, prepositional phrases). These words and phrases are in French probably due to the personal preference of the writer as the codeswitch does not belong into the other categories, like cultural switches. It is possible that the writer prefers using the French form instead of English for those words and phrases as he/she uses them only in French form in everyday life. It might also be easier for the writer to express the meaning in French. In addition, Lexander (2012, 153) states in her study that "[t]he informants confirm that they usually find French more romantic than other languages. According to Rama and other informants, 'I love you' cannot be expressed in Wolof with the same signification as *je t'aime*", which shows that language attitudes and feelings of language users towards a language can also have an impact on language choice. In addition, speakers/writers tend to choose the forms that are shorter and simpler and require the least effort instead of complicated and intricate forms (Halim and Maros 2014, 132). Similarly, Gumperz (1982) notes that shorter and easier forms are preferred by bilingual speakers. For example, the French codeswitch *chez toi* is shorter than the English equivalent 'at your place/home'.

Free switching category codeswitches include for example, *déductions* ('deductions'), *étranger* ('outsider'), *pas de commentaire* ('no comment'), *tapis roulant* ('conveyor belt'), *crédits* ('credits') and *consummation électrique* ('electricity consumption'). In many cases, the only difference between English and French word is only the "e" letter having an accent, but the writers still use the French form. Especially in written texts, the writers have more time to think about what they write and how and edit their text as writing is more planned discourse than speaking. The use of French words in these cases is probably a conscious choice, unless the writer automatically writes these words in French without thinking about it since it is a part of their idiolect.

5.2.3 Language discussion

Language discussion codeswitches were quite frequent in the data compared to many of the other categories as 31 codeswitches belonged to language discussion category making it the third most frequent function category. These involved discussion about the meaning of different French words or phrases, which form of the French word or phrase is correct in a specific context, providing clarification for other users about certain French words or discussion about French grammar. In addition, politeness of phrases was also discussed (*je m'excuse* vs *excusez-moi* vs *je suis navre*).

- (7) **papera** = paperwork, to those of us who write in English. Not that I mind French but I prefer one or the other. Let the insults fly.....
(User 6)

In example 7, the writer specifically clarifies the word *papera* for people who do not speak French. However, the following example from a different user is from the same discussion thread:

- (8) To any person who speaks and understands French there is a world of difference between paperwork and **papera** and if I was speaking to another French speaking English person about the totally over the top obsession with paper, boxes to tick and the evangelical desire to overcomplicate things I would never use the word paperwork (User 3).

In example 8, the user discusses differences between the English word paperwork and French word *papera* which literally translated means ‘paperwork’, but as the user points out the French word has a different connotative meaning. Similarly, in example 9, the writer describes the meaning of the French word *sympa* and how it is hard to translate precisely into English, *The Cambridge Dictionary* defines the French word *sympa* as ‘nice’ or ‘friendly’ (Cambridge Dictionary, s.v. “sympa,” adj.) The writer also provides a real-life example of how it is used from his past experiences.

- (9) **Sympa** is also hard to translate especially as you can use **sympa** to describe both people and places. A French person, to whom I gave a link to look up my village, came back and told me that my village was **très sympa**[:D] (User 4).

In example 10, the writer codeswitches for a different reason than in other language discussion cases:

- (10) I speak and write franglais, being in the bilingual household I am in, we speak franglais all the time. Just because you don't like it, is your

problem, as I'll not stop. If people don't want to read what I have said, when they see I've posted, they'll skip that bit of the post, **c'est tout, ca m'est egale**, or is that **je m'en fiche**[blink], that's life. I like and respect a lot of people on here, and I don't expect to agree with everyone, or please everyone either. I have my say, they have theirs, it usually works rather well without 'having a go at someone' (User 4).

The post is a direct answer to a post in example 9 which clarified the word *paperasse*. The writer specifically codeswitches to French because he/she wants to prove a point to the previous user that using French is okay and he/she has every right to use it in the forum. “C’est tout, ca m’est egale” means ‘that's all, it doesn't matter to me’ and “je m’en fiche” means ‘I do not care’. Clearly the French language is important to the writer and that is one reason for codeswitching. In addition, the writer states that they grew up in a bilingual household, so it is to be expected that the writer uses both languages.

Metalinguistic commentary automatically includes codeswitching as people discuss the meanings of foreign language words (Halmari 1997, 41) and similarly, discussion about the meaning and structure of words without examples from the given language would be burdensome (Kortelainen and Kolehmainen 2022, 36). By providing examples, clarifications and translations for words being discussed more people will be able to understand the discussion and the post becomes more comprehensible.

5.2.4 Quotations

Quotations was the fourth most frequent function category with approximately twenty switches. Quotations belonged to intersentential switches. Passages that contain codeswitched quotations or reported speech are easily identified (Gumperz 1982, 75–76). Quotations are commonly used when one reports what someone else has said or written. Reporting the content in the source language can be easier since the text is readable in its original language. Furthermore, there is no risk of meanings being altered or lost with translation. Codeswitched passages also retain the authenticity of the message (Halim and Maros 2014, 129). In addition, in codeswitched passages they can also act as a contextualisation cue that marks “voices” in conversation or written texts (Frick and Riionheimo 2013, 567). In the data, the codeswitches in this category were either extracts of textcopy-pasted from other websites (for example, a news article or a speech) or reported speech. The length of the quotations varied with some being quite short (a phrase or a sentence) and some were several paragraphs (for example, song lyrics copy-pasted to a post).

Extracts of text directly copy-pasted from other websites can be viewed as codeswitching specific to computer mediated communication (Kortelainen and Kolehmainen 2022, 30) as it makes use of available tools in computer-mediated communication (websites) that are not as easily available in spoken interactions:

- (11) I have dual nationality While I am on French soil I am considered as 'French' *un Français binational ou plurinational ne peut pas faire prévaloir sa ou ses autres nationalités lorsqu'il réside en France.* <https://www.service-public.fr/particuliers/vosdroits/F334> So I voted in a French election as a French national on the issues I have posted images of (User 7).

In example 11, the writer has copy-pasted one sentence from an official French website in order to support his statement about nationality. The writer has written the quotation in italics in order to separate it from his/her own text. In addition, the writer has provided a source for the quotation so other users can check themselves that the information is correct. Direct quotations in my data usually had the following functions in the data: supporting the user's argument or statement, acting as a conversation starter (user asks something regarding the quotation) or instructions (for example, how to find specific information on a website).

Reported speech occurred considerably fewer occurrences than direct quotations in my data as there were under ten occurrences of reported speech. Reported speech is about reporting what the writer/speaker themselves or someone else has said previously. Codeswitching in reported speech can serve the function of preserving the source language of what has been said or written (Lantto and Kolehmainen 2017). The instances of reported speech in my data involve reporting what the user themselves has said or the writer reports what he/she has heard someone else saying at some point earlier in time. For example:

- (12) As for bureaucrats, ours is not to wonder why. My French neighbour used to answer my frustrated comments with '**Mais, c'est comme ça**' (User 8).

In example 12, the user reports what his/her neighbour has said ("Mais, c'est comme ça" 'but that's how it is'), at some unspecified point in time in original language. The speech of another person is also separated from the rest of the text with quotation marks.

5.2.5 Humour

Some codeswitches could be classified as humour. At least a handful of forum users are from England, and as the discussion forum focuses on French culture, society and living in France,

it is to be expected that people joke about cultural differences when discussing cuisine for instance. France and England have been rivals in history several times. The mockery of British culture by the French and vice versa is quite common. A few humorous codeswitches in the data were about discussing English and French cuisine. This can be observed in example 13:

- (13) When my French friends ask me what I would give them to eat that was typically Brit, I always say, **mais le rosbif, bien sûr!** (User 9).

The *Oxford English Dictionary* defines the noun *Rosbif* in the following way: “[a]mong French-speakers: (a pejorative term for) an Englishman” (OED, s.v. “Rosbif,” n.). It is considered to be a well-known standard insult to the British or playful teasing and generally refers to the opinion that English cuisine is not that noteworthy when compared to French cuisine which is quite popular and known in the world. This can be observed also in other discussion threads in the discussion forum, but it also appears in other websites.

- (14) I'm a relatively new poster and don't find a bit of franglais in the slightest bit off-putting. In fact, it's quite inspiring because I just had to go and find the word for 'crumbs' with which to start this post. Still, each to their own. I'm off to watch Mapp & Lucia now, so **'au reservoir** (User 10).

In example 14, *au reservoir* is a humorous alteration of *au revoir* (‘goodbye’) after *reservoir* (‘storehouse’). It is considered a malaprop. Malapropism is defined as “[t]he ludicrous misuse of words, esp. in mistaking a word for another resembling it” (OED, s.v. “malapropism,” n.). The codeswitch originated from a Mapp and Lucia comic novel. In this case, one would have to be familiar with the novel in order to understand the meaning of the codeswitch.

Finally, the French codeswitch *Dunkerque* is categorised as a humorous given the surrounding context and that the writer writes “LOL” after it. *Dunkerque* is the French name for the war film ‘Dunkirk’ and the writer discusses the annoying habit that in France film names are often translated, which is unnecessary in the writer’s opinion.

5.2.6 Greetings/farewells

The frequency of *greetings/farewells* was quite rare as under 5 codeswitches could be categorised as a greeting or farewell. Greetings and farewells tend to be one of the first things

learned when moving to new linguistic surroundings as they are basic expressions that occur frequently and they occur at the start and at the end of a comment (Kortelainen and Kolehmainen 2022, 32). In my data, one informal greeting in the data occurs at the beginning of the post and greets another user: *Salut, Mint* ('Hello, Mint'). Another codeswitch that is a farewell occurs at the end of the post: *Amicalement* ('best wishes'). *Amicalement* is a typical informal farewell that can often be used in letters or emails. One greeting also occurs in the middle of the post:

- (15) So the point of your post is...you can vote. Well done. **Salut** and sing the **Marseillaise** and go and play patanque. I hope you have a proper curly French moustache BTW. LOL [User 1].

Example 15 has a humorous and mocking tone in response to a post made by another user, and it has several things that reference Frenchness. First, "Salut" is a French greeting, although here it also has a humorous tone. In addition, the codeswitch is meant to add to the Frenchness of the post. "Marseillaise" refers to *La Marseillaise*, which is the national anthem of France. Finally, "patanque" (spelled incorrectly in the post as it is spelled *pétanque* in English and French) which is a sport that originally emerged in France in the 19th century, although it is not considered a codeswitch, but a loanword.

5.3 Discussion

Only 150 out of 582 posts contained codeswitching, which suggests that codeswitching is not very frequent in the discussion forum overall when compared to the number of posts only in one language. However, 150 is still a considerable number of codeswitches, and in specific discussion threads the frequency of codeswitching was quite high, especially in threads where the discussion was about cuisine. Some codeswitches were used very often in a discussion threads. For example, the noun *gîte* occurred 10 times in different forms in one discussion thread. While codeswitching does take place, many probably prefer writing only in English as it might be easier and quicker than writing in French,

It is not surprising that intrasentential switching is much more frequent in the data since similar results have appeared also in other studies that focused on written codeswitching, such as Kortelainen and Kolehmainen (2022) who also analysed a discussion forum, and Yusuf and Fata (2018) who focused on codeswitching in a novel. Switching between sentences only occurred when reporting what someone else had said. Many users of the forum likely have some knowledge of French based on the amount of French used in the forum, but they are not

necessarily fluent speakers or writers and writing a complete sentence on their own in French can be demanding. Intrasentential codeswitching seems to be more effortless for the forum users since it can take place in any context by just using one word from the other language (greeting or discourse marker or some noun, which can fit into any context). Intersentential switching occurs less frequently as it mostly consisted of quotations or excerpts of reported speech, which did not occur frequently in the data. Khalil and Firda (2018, 34) argue in their study about codeswitching in parliamentary debates that intersentential codeswitching is used due to lack of equivalent terminology, mood of the speaker and to emphasize a point. In my study, intersentential codeswitching is also often used for these reasons as quoting an official and original source makes the argument stronger and convincing.

Quotations or excerpts of reported speech occurred in specific contexts when there was a need to cite something (for example, what someone has said or written in the news or citing official information). Another reason is that the main language of the discussion is English since majority of the posts are in English and it would be odd if the topic creator begins the thread by asking a question in English and another user answers mostly in French, unless there is a specific need for it (for example, providing a quotation to support an argument). In addition, the forum consists of users from around the world with different language skills, posts mainly in French would make the discussion thread less accessible for many forum users who are not fluent in French, and it would limit the number of people who could be part of the conversation in the discussion thread.

The grammar of French nouns could vary greatly. French nouns often received the English indefinite or definite article, but the noun itself follows French grammar by appearing in the gender (masculine or feminine) it has in French. In addition, French nouns only had the -s plural form in the data even though French grammar also has other forms for plural. Finally, the capitalisation patterns of French nouns varied greatly in the discussion forum. Some nouns were capitalised when they should not have been and some French nouns uncapitalised, but they should have been capitalised. Generally, the capitalisation depended on the writer's preference as in many cases the French nouns did not follow either the English or the French capitalisation rules. The variation in grammar might be due to informal nature of the forum as grammar is evidently not considered to be very important as long as the message itself is understandable and people might write quickly and not edit or proofread their posts.

The significant number of cultural codeswitches in the data is not surprising as Kortelainen and Kolehmainen (2022, 26) also noted in their study that “[a] great share of the Finnish codeswitches was culture-related” and the discussion forum is about anything to do with France and French culture. Furthermore, the discussion is also meant for people living in France or thinking about moving to France, so it is to be expected that cultural terms are mentioned often since many people likely have questions about living in France and where to go or who to talk to if one has a problem (for example, how to get a health insurance card). In addition, there are specific topic categories meant for discussing moving to France or talking about French culture (for example, food, language and history) in the discussion forum, but there are plenty of cultural switches in other topic categories as well (for example, *General Discussion*).

The frequency of *greetings/farewells* was also rare. This differs from the results in Kortelainen and Kolehmainen’s study (2022, 32) who note that greetings and farewells were typical sites of codeswitching in their data. It should be noted that greetings and farewells are rare even in English in most of the posts as the writers immediately start discussing their topic in discussion threads. It is possible that greetings and farewells are more frequent in other discussion threads that were not analysed for this study. In addition, the frequency of greetings and other salutations differs between discussion forums.

Another fact that should be noted is that cultural switches are very rarely translated or explained in English. Translating certain words and expressions may reduce their semantical accuracy, meaning that it is easier to render the term in its original language than to try to translate it accurately (Halim and Maros 2014, 131). Explanations in English only occur in *language discussion* category switches since for example, it involves discussion about the different meanings and connotations of certain French words or phrases. The writers on the forum expect adequate French language knowledge and also knowledge about French society and culture and concepts and how they work and writers also “can rely on this shared knowledge when conveying cultural meanings” (Kortelainen and Kolehmainen 2022, 29). The lack of translations and explanations can cause confusion for some people if they do not know the context of the discussion thread and do not know much about French society. Several cultural switches are very context dependant as they can have multiple meanings (for example, *CV*), or the meaning is not literal (Mutuelle ‘mutual’ vs ‘health insurance’).

However, if one is aware of the context of the discussion thread and has knowledge about the different acronyms and French concepts then it should not be a problem to understand the discussion and which meaning of the word is correct.

Kortelainen and Kolehmainen (2022, 36) note that “Most of the metalinguistic commentaries were part of the forum’s discussion category “Finnische Sprache” (‘Finnish language’), and they accordingly occurred in discussions on the properties of Finnish language [...]”.

However, in my data most of the *language discussion* category switches came from *French culture* and *Legal and Finance* topic categories, which suggests that metalinguistic commentary can occur quite frequently in other topic categories as well. The topic category *French language* was not included in this study, but it is likely that there are many codeswitches in that topic category.

6 Conclusion

The objective of this study was to investigate codeswitching between English and French in the online discussion forum *CompleteFrance* where people can ask for advice and discuss anything about moving to France or French culture and society. The main language on the forum was English, but the users often codeswitched to French in different discussion threads in different contexts for various purposes.

308 codeswitches in total were identified across 12 different discussion threads which indicates that codeswitching takes place in different contexts. However, while 582 posts in total were included in this study, only 150 posts contained codeswitching. The structural analysis showed that intrasentential codeswitching is used much more frequently in the discussion forum than intersentential codeswitching as 82% of the French codeswitches analysed for this study were intrasentential. Intrasentential codeswitches consisted of many different French words and phrases while intersentential (8%) switches consisted of quotations that could vary in length from short phrases to several paragraphs. The remaining codeswitches (10%) were considered metalinguistic commentary. These involved discussion about various aspects of French language. For example, the meaning of certain French words and how to translate them was discussed. In addition, discussion about formality also took place in one thread. Most of the French codeswitches were common nouns, proper nouns or noun phrases as they often referred to concepts and things specific to French culture and society or to French places, organisations and institutions. Other parts of speech and types of phrases (for example, adjectives, adverbs and interjections) were much rarer in the data as each one had around ten occurrences or even less in the data. Formulaic expressions were also included in the structural analysis. Formulaic expressions are multiword units or phrases. They are generally well-known phrases that are often used and also many of them are known outside French. In the data, many of them involved wishing luck to someone or welcoming someone.

The grammar of the French nouns had variation in terms of capitalisation, plural forms, articles and gender. Incorrect grammatical forms appeared at times, which is probably due to the informal nature of the forum. A few of hybrid forms also appeared where both English and French grammar elements were visible (*dépends*).

The functional analysis showed that the functions of the French codeswitches correspond to the various codeswitching function categories that have also appeared in previous research. These include *cultural switches*, *quotations*, *greetings/farewells* and *humour*. In addition, *free switching* and *language discussion* categories were used for this study. The frequency of cultural switches was especially high as there were much more cultural switches than any other codeswitches. Cultural switches were divided to three thematic categories (society, cuisine and places) and they referred to concepts and objects specific to French culture and society. Halim and Maros (2014, 133) note in their conclusion that codeswitching “is also used as a strategy to communicate and negotiate meanings effectively in their electronic discourse”, which applies to many of the cultural switches in my study. The high frequency of cultural switches was expected considering the purpose of the discussion forum from which the data was collected since it is about discussing anything to do with France. Furthermore, cultural switches were often used because there was no English equivalent word available that conveys the exactly the same meaning. In addition, cultural switches could function as contextualisation cues by defining a concept as French, which was also noted in Kortelainen and Kolehmainen’s study (2022, 37) with Finnish cultural switches and cultural switches in their study also had a high frequency.

One feature specific to computer-mediated communication is quotations. Codeswitches that belonged to the quotations category consisted mostly of direct quotations copy-pasted from other sites which could vary in length and there were also a few instances of reported speech. Codeswitches that belonged to humour category were often about the rivalry between English and French and especially English cuisine was discussed humorously. *Greetings and farewells* were quite rare in the data as the writers on the forum rarely used greetings or farewells in their posts which is not surprising since greetings and farewells probably appear more in emails and text messages instead of discussion forums, but it can also depend on the formality of the platform used for communication.

Language discussion codeswitches were also quite frequent in the data. These involved discussion about the French language and how to use it. The discussion was about the meanings of various French words and phrases, what form of the word is correct in a specific context, politeness or how to translate a word or phrase that can not be translated literally to English. Finally, free switching codeswitches consisted of French single words or phrases that did not fit into any of the function categories. The codeswitches were likely used because the writer preferred to use the French form of the word or phrase. That might be due to the

personal preference of the writer, or the writer is used to using the French form of particular words or some other unknown reason. In many cases, the words had minor differences between English and French.

My results complement the analysis of the structure and functions of codeswitching, specifically between English and French languages in one discussion forum. The findings in this study account for only the discussion forum analysed for this thesis. Analysing a different discussion forum could provide different patterns of codeswitching (For example, more greetings). Similarly, only a part of the discussion forum was analysed. Analysing different discussion threads for a new study bigger in scope could be useful. Furthermore, analysing a different language pair could provide different results. For example, French and Italian are much similar than English and French which would likely show in analysis. My study focused on the written medium of codeswitching. It would be interesting to know if the users who often codeswitched in their posts also codeswitch as often and in the same way in spoken communication. In addition, linguistic background and language attitudes can have a role in how often a person codeswitches. A study where the participants are also interviewed about their linguistic backgrounds in addition to analysing their written texts would be useful in order to better understand for what reason and when people codeswitch. Web forums continue being an important environment for producing text. Many different topics can be studied by analysing discussion forums, such as multilingualism, and many research opportunities still remain.

References

- Auer, Peter. 1984. *Bilingual Conversation*. Amsterdam: John Benjamins Publishing Company.
- Auer, Peter. 1999. "From Codeswitching via Language Mixing to Fused Lects: Toward a Dynamic Typology of Bilingual Speech." *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior* 3, no. 4: 309–332
- Bailey, Benjamin 2007. Multilingual forms of talk and identity work. *Handbook of Multilingualism and Multilingual Communication*, eds. Peter Auer & Li Wei. 341–369. Berlin: Mouton de Gruyter
- Barančicová, Jana, and Jana Zerzová. 2015. "English as a Lingua Franca Used at International Meetings." *Journal of Language and Cultural Education* 3, no. 3: 30–51.
- Cambridge University Press. (n.d.). Sympa. In *Cambridge Dictionary*. Retrieved March 20, 2024
- Eastman, Carol and Monica Heller. 1990. Codeswitching: Anthropological and Sociolinguistic Perspectives (Contributions to the Sociology of Language 48). Berlin: Mouton de Gruyter, 1988. Pp. 278. *Language in Society*, 19(3), 442–447. <https://doi.org/10.1017/S004740450001469X>.
- Frick, Maria and Helka Riionheimo. 2013. Bilingual voicing: A study of code-switching in the reported speech of Finnish immigrants in Estonia. *Multilingua*. 32. 10.1515/multi-2013-0028.
- Gardner-Chloros, Penelope. 2009. *Code-Switching*. Cambridge: Cambridge University Press.
- Graddol, David. 2000. The future of English? A guide to forecasting the popularity of the English language in the 21st Century. London: British Council. Accessed 24 October 2022. <https://www.teachingenglish.org.uk/article/future-english>
- Gumperz, John J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press
- Halim, Nur Syazwani and Marlyna Maros. 2014. "The Functions of Code -switching in Facebook Interactions." *Procedia - Social and Behavioral Sciences* 118: 126-133.
- Halmari, Helena 1997. *Government and Codeswitching: Explaining American Finnish*. Amsterdam: Benjamins. <https://doi.org/10.1075/sibil.12>
- Hinrichs, Lars. 2006. *Code-switching on the Web: English and Jamaican Creole in E-mail Communication*. Amsterdam: John Benjamins

- Khalil, Shafiyah and Firda Mohammad. 2018. Inter-Sentential and Intra-Sentential Code Switching in Parliamentary Debate. *International Journal of Modern Languages And Applied Linguistics*. 2. 29. 10.24191/ijmal.v2i4.7691.
- Kortelainen, Henna, and Leena Kolehmainen. 2022. “‘Gruß Aus Saksa’: Multilingual Practices in a German Expatriate Online Community in Finland.” *Neuphilologische Mitteilungen* 123, no. 1: 6–48.
- Kytölä, Samu. 2012. “Multilingual Web Discussion Forums: Theoretical, Practical and Methodological Issues.” In *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*, 106–127. London: Routledge.
- Kuźelewska, Elżbieta. 2020. “Quo Vadis English? The Post-Brexit Position of English as a Working Language of the EU.” *International journal for the semiotics of law = Revue internationale de sémiotique juridique* 34, no. 5: 1417–1432.
- Lantto, Hanna, and Leena Kolehmainen. 2017. “Basque-Spanish Bilinguals and Reported Speech. Translation and Code-Switching in the Basque Context of Language Revitalisation.” *Trans-Kom: Journal of Translation and Technical Communication* 10, no. 2: 214–241.
- Lexander, Kristin Vold. 2012. “Analyzing Multilingual Texting in Senegal: An Approach for the Study of Mixed-Language SMS.” In *Language Mixing and Code-Switching in Writing*, 155–178. London: Routledge.
- Myers-Scotton, Carol. 1993. *Duelling Languages*. Oxford: Clarendon Press
- Myers-Scotton, Carol. 1998. “A Theoretical Introduction to the Markedness Model.” In *Codes and Consequences*. New York, NY: Oxford University Press.
- Myers-Scotton, Carol. 2011. *Multiple voices: an introduction to bilingualism*. Oxford: Blackwell Publishing
- OED: *Oxford English Dictionary*. Oxford: Oxford University Press. Accessed 28 March 2024. <https://www.oed.com/>
- Paolillo, John. 1999. The virtual speech community: Social network and language variation on IRC. *Journal of Computer Mediated Communication*, 4/4.
- Rajala, Joonas. 2019. *Code-Switching Between English and French in The Montreal Gazette*. Accessed 23 November 2022. <https://www.utupub.fi/handle/10024/148648>
- Raukko, Jarno. 2017. Names in contact: Linguistic and social factors affecting exonyms and translated names. in T Ainiala & J-O Östman (eds), *Socio-onomastics: The pragmatics of names. Pragmatics & Beyond New Series*, vol. 275, John Benjamins, Amsterdam, 93–125.

- Sebba, Mark and Tony Wooton. 1998. We, they and identity: sequential versus identity-related explanation in code-switching. *Code-switching in Conversation: Language, Interaction and Identity*, ed. Peter Auer. 460–500. London: Routledge.
- Wei, Li 1998. The ‘why’ and ‘how’ questions in the analysis of conversational code-switching. *Code-switching in Conversation: Language, Interaction and Identity*, ed. Peter Auer. 280–316. London: Routledge.
- WorldData. 2023. English speaking countries. Accessed 20 March 2023.
<https://www.worlddata.info/languages/english.php>
- Yankova, Diana and Irena Vassileva. 2015. “Functions and Mechanisms of Code-Switching in Bulgarian Canadians.” *Études canadiennes*: 103–121.
- Yusuf, Yunisrina, Ika Fata and Cynthia. 2018. Types of Indonesian-English code-switching employed in a novel. *Kasetsart Journal of Social Sciences*. 10.1016/j.kjss.2018.02.004.

Appendices

Finnish Summary

Tämä tutkielma tarkastelee koodinvaihtoa englannin ja ranskan kielten välillä *CompleteFrance*-nimisellä internetin keskustelufoorumilla, jossa voi keskustella muun muassa ranskalaisesta kulttuurista ja kirjallisuudesta sekä ranskan kielestä. Sekä englantia että ranskaa puhutaan laajalti eri maanosissa, mutta englannin osaajia on huomattavasti enemmän kuin ranskaa osaavia. Lisäksi englannin kieli toimii *lingua francana* monelle ihmiselle ja sitä käytetään monissa eri akateemisissa ja ammatillisissa piireissä. Tämän tutkielman tavoitteena on selvittää, missä yhteyksissä englannin ja ranskan kielten välistä koodinvaihtoa tapahtuu ja missä yhteyksissä sitä tapahtuu eniten, millaista koodinvaihtoa foorumilla esiintyy, ja mitkä ovat koodinvaihdon funktiot. Aihe on oleellinen, koska koodinvaihtoa englannin ja ranskan kielten välillä ei ole tutkittu paljon. Englanti ja ranska ovat hyvä kielipari tähän tutkimukseen, sillä englanti on globaali kieli ja hyvin tärkeässä asemassa viestinnässä eri aloilla maailmalla ja ranska on myös laajalti puhuttu kieli.

Koodinvaihto määritellään yleisesti kahden tai useamman kielen käyttönä samassa keskustelussa (Eastman ja Heller 1988, 1). Tämä keskustelu voi olla kirjallinen tai suullinen. Koodinvaihto käsitteenä on ajan myötä saanut hieman toisistaan poikkeavia määritelmiä sekä kilpailevia termejä. Yleensä määritelmä riippuu siitä, mistä näkökulmasta aihetta tutkitaan ja millä menetelmillä aihetta lähestytään. Koodinvaihtoa voi tapahtuma myös silloin, kun puhutaan vain yhtä kieltä, jos puhuja tai kirjoittaja käyttää esimerkiksi eri murteita keskustelussa. Yksi tutkimuksissa usein käytetty teoria koodinvaihdosta on Myers-Scottonin vuonna 1993 kehittämä teoria, joka käsittelee klassista koodinvaihtoa. Sen mukaan yksi koodinvaihdon kielistä on hallitseva (*Matrix Language*) ja toinen alisteinen (*Embedded Language*) (Myers-Scotton 1993, 3–4). Tässä tutkielmassa englannin kieli toimii hallitsevana kielenä ja ranska on sille alisteinen. Jotkut tutkijat erottavat toisistaan myös lauseiden ja lausekkeiden välillä tapahtuvan koodinvaihdon (*intersentential codeswitching*) ja lauseen tai lausekkeen rajojen sisällä tapahtuvan koodinvaihdon (*intrasentential codeswitching*) (Kortelainen and Kolehmainen 2022, 11–12).

Sen erottaminen, onko yksittäinen sana koodinvaihtoa vai lainasana toisesta kielestä, voi olla haasteellista, koska koodinvaihdolla ja lainaamisella on yhteisiä piirteitä. Lainasanat ovat osa sen kielen kieliopillista järjestelmää, johon ne on lainattu, ja lainasanat omaksuvat uuden

kielen morfologiset piirteet ja siirtyvät kielen syntaktisiin rakenteisiin (Gumperz 1982, 66). Jos vierasperäinen sana esiintyy *Oxford English Dictionary* -sanakirjassa, se luokitellaan lainsanaksi, eikä koodinvaihdoksi, ellei OED:ssä erikseen mainita, että sana on nimenomaan ranskan kielelle ominainen tai vanhentunut eli sitä ei käytetä enää nykyään englannin kielessä.

Koodinvaihdon syitä on myös tutkittu laajasti. Koodinvaihto voi toimia myös kontekstivihjeinä, joiden avulla kuulijat ja puhujat voivat tulkita, millaisesta toiminnasta on kyse, miten semanttista sisältöä pitäisi tulkita ja miten kukin lause liittyy sitä edeltävään tai seuraavaan lauseeseen (Gumperz 1982, 131). Ihmiset käyttävät ja havaitsevat kontekstivihjeitä usein, mutta yleensä emme puhu niistä suoraan ja havaitsemme ne usein tiedostamattomasti (ibid.).

Koodinvaihto liittyy myös identiteettiin, joka usein ilmaistaan termeillä me-koodi ja he-koodi (Gumperz 1982). Vähemmistöryhmän käyttämä kieli toimii me-koodina ja se on tarkoitettu ryhmän sisäiseen ja epäviralliseen toimintaan, kun taas he-koodi on enemmistöryhmän käyttämä kieli ja se liittyy viralliseen, ryhmän ulkopuoliseen toimintaan (Gardner-Chloros 2009, 56). Me-koodin ja he-koodin asema ei ole vakaa (Sebba 1998), ja identiteetti ja identiteettiyhdistelmät rakentuvat vuorovaikutuksen aikana sen sijaan, että ne olisivat ennalta määrättyjä (Bailey 2007).

Koodinvaihdon tutkimiseen kirjallisessa aineistossa liittyy useita haasteita. Yksi syy on se, ettei ole olemassa teoreettisia viitekehyksiä, jotka olisi suunniteltu erityisesti koodinvaihdon ja kielivalintojen tutkimiseen digitaalisessa mediassa (Lexander 2012, 146). Monet tutkijat ovat soveltaneet teorioita, jotka on alun perin tarkoitettu puhutun kielen tutkimiseen, mutta nämä teoriat eivät täysin ota huomioon kirjoitettujen tekstien monikielisyyttä (Hinrichs 2006). Silti ne tarjoavat hyödyllisiä tietoja koodinvaihdon tutkimiseen. Verkkofoorumin kieltä ei pidetä puhuttuna kielenä, sillä verkkofoorumeilla ei ole tyypillisesti ääntä (ibid.). Verkkokieli on kirjoitettua ja asynkronista, mutta se voi vaihdella spontaanista tai tiedostamattomasta erittäin tyyliteltyyn, huolellisesti rakennettuun kieleen (ibid.). Verkkofoorumin käyttäjät päättävät, millaista kieltä he käyttävät ja muokkaavat foorumia omiin tarpeisiinsa (ibid.). Keskustelufoorumeilla on myös usein joku tietty tarkoitus ja kohderyhmä. Vaikka keskustelupalstojen kielellinen aineisto on helposti kaikille saatavilla sekä tallennettavissa ja arkistoitavissa tutkimusta varten, tutkijan pitää tehdä metodologisia päätöksiä perehtymisen ja kontekstualisoinnin asteen suhteen (Kytölä 2012, 110). Lisäksi tutkijan pitää tehdä eettisiä päätöksiä aineiston käytettävyydestä ja anonymisoinnista, kuten verkkofoorumin

käyttäjätunnusten poisjättäminen ja käyttäjien henkilötietojen poisto, jos niitä löytyy. Tässä tutkimuksessa käyttäjätunnukset anonymisoidaan pseudonymisoimalla.

Materiaalit tutkimusta varten on kerätty *CompleteFrance*-keskustelufoorumilta, jossa käyttäjät voivat keskustella ja pyytää apua erilaisissa Ranskaan liittyvissä asioissa (matkustaminen, muutto, palvelut, kulttuuri). Foorumin pääkielenä toimii englanti, sillä keskustelu tapahtuu enimmäkseen englanniksi, mutta monissa viesteissä esiintyy myös jokin muu kieli, useimmiten ranska. Kaikkea foorumin sisältöä voi tarkastella ilman, että tarvitsee luoda tiliä foorumille. Foorumin käyttäjät tulevat eri maista eri puolilta maailmaa, ja heillä on erilainen kielitaito, ja useimmat heistä käyttävät englantia foorumilla. Tästä syystä on huomattava, että tämän tutkimuksen aineisto ei ole peräisin pelkästään englannin kielen L1-puhujilta, joiden L2 olisi ranska. Tämä ei kuitenkaan ole ongelma, koska kunkin käyttäjän L1- ja L2-taidoilla ei ole merkitystä tutkimuksen kannalta. Tässä tutkimuksessa tarkastellaan vain niitä viestejä, jotka ovat pääosin englanniksi ja sisältävät koodinvaihtoa ranskaan, koska yksikieliset viestit ja muut kielet eivät ole merkityksellisiä tälle tutkimukselle. Foorumin viestien lisäksi myös keskustelupalstojen otsikot otetaan huomioon, jos ne sisältävät koodinvaihtoa.

Keskustelufoorumin kaikki viestit ovat julkisia ja avoimia kaikille. Tämän vuoksi ei tarvittu tutkimuslupia foorumin ylläpitäjiltä tai käyttäjiltä, joiden viestejä analysoitiin tutkimusta varten. Tässä tutkimuksessa käytetyissä esimerkeissä käyttäjien nimet ovat anonymisoitu. Lisäksi tutkimuksessa mainitut esimerkit ovat sävyltään neutraaleja, jotta ne eivät loukkaa ketään. Tutkimuksessa keskitytään vain foorumin viesteihin ja keskusteluketjujen otsikoihin, joten foorumin käyttäjien käyttäjätunnuksia ei oteta huomioon.

Aineisto kerättiin vuoden 2022 lokakuun ja vuoden 2023 elokuun välillä. Koodinvaihtoa sisältäviä viestejä kerättiin yhteensä 12 eri keskusteluketjusta. 582 viestiä analysoitiin perusteellisesti koodinvaihtojen varalta ja 150 viestiä sisälsi koodinvaihtoa. Koodinvaihtoa sisältävät viestit tallennettiin Word-tiedostoon. Suurin osa viesteistä sijoittuu vuosille 2014–2020. Jokainen koodinvaihto ranskaan kaikissa viesteissä lihavoitiin ja analysoitiin ottaen huomioon myös ympäröivän tekstin ja edelliset viestit. Myös aiemmat tutkimukset on otettu huomioon analysoinnissa. Koodinvaihtojen rakennetta analysoitiin ja koodinvaihdot jaettiin 3 eri luokkaan: lauseiden välillä tapahtuva koodinvaihto, lauseen rajojen sisällä tapahtuva koodinvaihto ja metalingvistinen keskustelu. Koodinvaihdot jaettiin myös sanaluokkien perusteella. Lopuksi analysoitiin vielä koodinvaihdon funktioita. Koodinvaihdon funktiot on

jaettu seuraaviin luokkiin: *kulttuurillinen, vapaa koodinvaihto, huumori, kielikeskustelu, lainaukset ja tervehdykset/hyvästit*.

Koodinvaihtoja ranskaan oli aineistossa yhteensä 305 kappaletta, joista 249 (82 %) oli lauseen tai lausekkeen rajojen sisällä tapahtuvaa koodinvaihtoa. Tähän luokkaan kuuluivat ranskankieliset sanat (esimerkiksi *salle*) ja lausekkeet (esimerkiksi substantiivilausekkeet), joiden pituus vaihteli. Lauseiden tai lausekkeiden välissä tapahtuvaa koodinvaihtoa aineistossa oli vain 25 kappaletta (8 %) ja ne koostuivat useimmiten lainauksista muilta sivustoilta. Lainauksien pituus vaihteli. Jotkut olivat lyhyitä (esimerkiksi vain yksi lause), kun taas jotkut lainaukset sisälsivät monta kappaletta pelkästään ranskaksi. Loput 31 koodinvaihtoa kuuluivat metalingvistisen keskustelun luokkaan. Useimmiten ne koostuivat erilaisista lauseista tai virkkeistä, joissa oli useita puheosia, eivätkä siten kuuluneet rakenneanalyysin piiriin. Metalingvistisen keskustelun luokka koostui enimmäkseen koodinvaihdosta, jotka sisälsivät keskustelua itse ranskan kielestä ja tiettyjen sanojen erilaisten merkitysten vertailusta englannin ja ranskan välillä. Näitä käsitellään enemmän funktioanalyysissä.

Koodinvaihdossa esiintyi eniten yleisnimiä, sillä niitä löytyi aineistosta yhteensä 131 kappaletta (54,1 %). Ne viittasivat usein kulttuurisidonnaisiin käsitteisiin, termeihin ja sanoihin, jotka mainitaan usein ja ovat yleisiä ranskalaisessa yhteiskunnassa. Näihin kuuluivat esimerkiksi ruokien nimet, paikat, ammatit ja ranskalaisen yhteiskunnan käsitteet. Yleisnimet viittasivat myös arkisiin esineisiin ranskaksi, kuten liukukäytävä tai kävelytie, jotka eivät ole ominaista ranskalaiselle kulttuurille. Myös ranskan kielen artikkeleita ja pronomineja esiintyi joskus, mutta ei yhtä usein kuin englannin kielen. Substantiivilausekkeita löytyi aineistosta 31 (12,8 %) kappaletta. Substantiivilausekkeet koostuivat enimmäkseen ranskankielisistä substantiiveista, joita edelsi esimerkiksi artikkeli tai adjektiivi. Tällaisia olivat esimerkiksi *repas anglais* ('englantilainen aterial'), *les filles* ('tytöt'), *un étranger* ('ulkomaalainen') ja *petite rue* ('pieni katu'). Ranskassa adjektiivit sijoittuvat usein substantiivin jälkeen, mikä voidaan havaita seuraavassa koodinvaihdossa: *Centres Commerciales* ('ostoskeskukset'). Yleisesti ottaen substantiivilausekkeet viittasivat myös yleisiin ja arkipäiväisiin sanoihin ja käsitteisiin.

Erisnimiä löytyi aineistosta yhteensä 43 (17,7 %) kappaletta. Ne koostuivat usein ilmauksista, jotka siirtyvät ilman muutoksia kielestä toiseen (Raukko 2017, 105–106). Erisnimet viittasivat usein aineistossa tiettyihin paikkoihin (*Côte d'Azur*), järjestöihin (*Médecins sans frontières*)

Lääkärit ilman rajoja') ja erilaisiin yhteiskunnallisiin ja sosiaalisiin järjestelmiin ranskassa. Jotkut koodinvaihdot viittasivat myös ranskalaisiin sanomalehtiin, nettisivustoihin tai ranskan poliittisiin puolueisiin. Eri lyhenteiden merkitystä ei ole foorumilla selitetty tai käännetty.

Muut rakenteelliset luokat esiintyivät harvemmin. 13 (4,5 %) koodinvaihtoa luokiteltiin adjektiiveiksi. Adverbeja oli aineistossa vain 5 (2 %) kappaletta. Lisäksi prepositiolausekkeita ja interjektioita oli vain 4 kappaletta molempia (1,7 %). Verbejä löytyi 2 kappaletta (0,8 %). Vakiintuneita ilmauksia oli 11 kappaletta (4,5 %). Ne ovat erillään muista rakenteista, kuten substantiiveista ja adjektiiveista. Vakiintuneita ilmauksia ovat esimerkiksi *bonne chance* tai *bon courage* ('hyvää matkaa' tai 'tsemppiä') tai *c'est la vie* ('sellaista elämä on').

Koodinvaihdolla oli myös monia eri funktioita. Enemmistöllä koodinvaihdosta oli jokin kulttuurillinen funktio, sillä yli 150 koodinvaihtoa kuului tähän funktioluokkaan. Kaikkiin muihin funktioluokkiin kuuluu alle 50 koodinvaihtoa. Erityisesti erisnimillä ja substantiivilausekkeilla oli kulttuurillinen funktio. Kulttuurikoodinvaihdot viittasivat usein käsitteisiin ja esineisiin, jotka ovat ominaisia ranskalaiselle kulttuurille.

Kulttuurikoodinvaihdot jaettiin myös kolmeen eri teemaluokkaan: *ruoka*, *yhteiskunta* ja *paikat*. Suurin teemaluokka oli *yhteiskunta*. Tähän teemaluokkaan kuuluvat koodinvaihdot viittasivat esimerkiksi terveydenhuoltoon, instituutioihin, sosiaaliturvaan ja ammatteihin. Toiseksi suurin luokka oli *ruoka*. Tähän luokkaan kuuluvat koodinvaihdot viittasivat erilaisiin ruokiin, viineihin ja yleisesti ruoanlaittoon. *Paikka*-teemaluokan koodinvaihdot viittasivat erilaisiin puistoihin, rakennuksiin, kauppoihin ja maantieteellisiin paikkoihin.

Noin 35 koodinvaihtoa kuului *vapaa koodinvaihto* -funktioluokkaan. Nämä koostuivat erilaisista ranskankielisistä yksittäisistä sanoista tai lyhyistä lauseista. Nämä koodinvaihdot johtuivat todennäköisesti kirjoittajan omasta mieltymyksestä, sillä ne eivät sopineet mihinkään muuhun funktioluokkaan. On mahdollista, että kirjoittaja suosii enemmän ranskankielisiä muotoja tietyistä sanoista, koska hän on tottunut käyttämään ranskankielistä muotoa omassa elämässään. Monissa tapauksissa ainoa ero englannin- ja ranskankielisen sanan välillä oli vain e-kirjaimen päällä oleva aksentti, mutta kirjoittaja käytti silti ranskankielistä muotoa. Erityisesti kirjoitetuissa teksteissä kirjoittajilla on enemmän aikaa miettiä, mitä he kirjoittavat ja miten muokkaavat tekstiään. Ranskankielisten sanojen käyttö näissä tapauksissa on luultavasti tietoinen valinta, ellei kirjoittaja automaattisesti kirjoita näitä sanoja ranskaksi ajattelematta sitä, koska se on osa hänen idiolektiaan.

31 koodinvaihtoa kuului *kielikeskustelu*-funktioiluokkaan. Tähän luokkaan liittyi keskustelua erilaisten ranskan kielen sanojen tai fraasien merkityksestä tai siitä, mikä on sanan oikea muoto tietyssä asiayhteydessä. Lisäksi keskusteltiin ranskan kielen kieliopista sekä fraasien kohteliaisuudesta.

Noin 20 koodinvaihtoa kuului *lainaukset*-funktioiluokkaan. Sitaatteja/lainauksia käytetään yleisesti silloin, kun kerrotaan, mitä joku toinen on sanonut tai kirjoittanut. Sisällön raportointi lähdekielellä voi olla helpompaa, koska tekstin voi lukea alkuperäiskielellä. Lisäksi ei ole vaaraa, että merkitykset muuttuvat tai katoavat käännöksessä. Aineistossa tähän luokkaan kuuluvat koodinvaihdot olivat joko otteita muilta verkkosivuilta (esimerkiksi uutisartikkeli tai puhe) tai raportoitua puhetta. Lainauksien pituus vaihteli siten, että osa niistä oli melko lyhyitä (lause tai virke) ja osa useamman kappaleen mittaisia (esimerkiksi laulun sanoitukset). Raportoitua puhetta esiintyi aineistossa huomattavasti vähemmän kuin suoria lainauksia, sillä sitä esiintyi alle 10 kertaa. Raportoidussa puheessa on kyse siitä, että raportoidaan sitä, mitä kirjoittaja/puhuja itse tai joku muu on aiemmin sanonut. Koodinvaihto raportoidussa puheessa voi palvella puhutun tai kirjoitetun lähdekielen säilyttämistä (Lantto ja Kolehmainen 2017).

Osa koodinvaihdoista luokiteltiin *huumori*-funktioiluokkaan. Ainakin jotkut foorumin käyttäjistä olivat kotoisin Englannista, ja koska keskustelupalstalla keskitytään ranskalaiseen kulttuuriin ja yhteiskuntaan sekä Ranskassa elämiseen, on odotettavissa, että ihmiset vitsailevat kulttuurieroista keskusteltaessa esimerkiksi ruoasta. On yleistä, että ranskalaiset vitsailevat englantilaisesta kulttuurista ja päinvastoin. Muutama humoristinen koodinvaihto aineistossa koski keskustelua englantilaisesta ja ranskalaisesta ruoasta.

Tervehdykset/hyvästelyt-luokkaan liittyvät koodinvaihdot olivat harvinaisia, sillä niitä löytyi alle 5 aineistosta. Tervehdyksiin ja hyvästelyyn liittyvät sanat opitaan yleensä nopeasti uudessa kieliympäristössä, sillä ne ovat perusilmaisuja (Kortelainen ja Kolehmainen 2022, 32).

Ei ole yllättävää, että lauseen tai lausekkeen rajojen sisällä tapahtuva koodinvaihto oli aineistossa yleisintä. Samankaltaisia tuloksia on ilmennyt myös muissa tutkimuksissa, joissa on keskitytty kirjalliseen koodinvaihtoon, kuten Kortelaisen ja Kolehmaisen (2022) tutkimuksessa, jossa myös analysoitiin keskustelufoorumia, sekä Yusufin ja Fatan (2018) tutkimuksessa, joka keskittyi romaanin koodinvaihtoon. Monilla foorumin käyttäjillä on todennäköisesti jonkin verran ranskan kielen taitoa foorumilla käytetyn ranskan kielen määrän

perusteella, mutta he eivät välttämättä ole sujuvia puhujia tai kirjoittajia. Lainauksia tai otteita raportoidusta puheesta käytettiin tietyissä yhteyksissä, kun oli tarve siteerata jotakin (esimerkiksi sitä, mitä joku on sanonut tai kirjoittanut uutisissa tai jotakin virallista tietoa). Foorumin pääkieli on englanti, koska suurin osa viesteistä on englanniksi, ja olisi erikoista, jos keskustelun luoja aloittaisi viestiketjun esittämällä kysymyksen englanniksi ja toinen käyttäjä vastaisi pääasiassa ranskaksi, ellei siihen ole erityistä tarvetta. Ranskan kielen substantiivien kielioppi vaihteli. Ranskankielisillä substantiiveilla oli aineistossa vain -s-pääte monikossa, vaikka ranskan kieliopissa on muitakin monikon muotoja. Osa substantiiveista oli myös kirjoitettu isolla alkukirjaimella, vaikkei olisi pitänyt, ja osalla ranskalaisista substantiiveista puuttui iso alkukirjain. Yleisesti ottaen isojen alkukirjainten kirjoittaminen riippui kirjoittajan mieltymyksestä, sillä monissa tapauksissa ranskankieliset substantiivit eivät noudattaneet englannin- tai ranskankielisiä isoja alkukirjaimia koskevia sääntöjä.

Kulttuuristen koodinvaihtojen merkittävä määrä aineistossa ei ole yllättävää, sillä myös Kortelainen ja Kolehmainen (2022, 26) totesivat tutkimuksessaan, että ”[suuri] osa suomalaisista koodinvaihdoista oli kulttuurisidonnaisia” ja keskustelupalstalla on kyse kaikesta Ranskaan ja ranskalaiseen kulttuuriin liittyvästä. Tervehdykset/hyvästelyt sen sijaan olivat harvinaisia. Tämä poikkeaa Kortelaisen ja Kolehmaisen (2022, 32) tutkimustuloksista, sillä he toteavat, että tervehdykset ja jäähyväiset olivat heidän aineistossaan tyypillisiä koodinvaihdon kohteita. Aineistossa tervehdykset ja jäähyväiset olivat useimmissa viesteissä harvinaisia jopa englanniksi, koska kirjoittajat aloittavat heti keskustelun omasta aiheestaan. On mahdollista, että tervehdykset ja hyvästelyt ovat yleisempiä muissa viesteissä, joita ei analysoitu tätä tutkimusta varten. Lisäksi tervehdysten ja hyvästelyjen yleisyys vaihtelee keskustelufoorumien välillä.

Kulttuurisia koodinvaihtoja käännettiin tai selitettiin englanniksi hyvin harvoin. Tiettyjen sanojen ja ilmausten kääntäminen saattaa vähentää niiden semanttista tarkkuutta, jolloin on helpompi esittää termi alkuperäiskielellä kuin yrittää kääntää se tarkasti (Halim ja Maros 2014, 131). Englanninkielisiä selityksiä esiintyi vain *kielikeskustelu*-funktioaluokassa.