

“I am a bilingual because I feel like one” – language identity
of Finnish IB-students

Laura Mäkelä

MA Thesis

English, Degree Programme for Language Teaching and Learning

School of Languages and Translation Studies

Faculty of Humanities

University of Turku

October 2020

UNIVERSITY OF TURKU

School of Languages and Translation Studies / Faculty of Humanities

MÄKELÄ, LAURA: "I am a bilingual because I feel like one" – language identity of Finnish IB-students

MA Thesis, 60 p., appendices 10 p.

English, Language Learning and Teaching Degree Programme

October 2020

This thesis examines the language identities of Finnish International Baccalaureate-students. In this study, the focus is on subjective experiences of bilingualism and the students' definitions of bilingualism. Very few studies have been conducted on psychological aspects of IB-students. Furthermore, previous studies on IB-students' language identities have concentrated on qualitative aspects. Thus, the main goal is to create a more generalized view of IB-students' language identity.

The thesis answers three questions: 1) what are the language identities of Finnish IB-students like, 2) to what extent do differences in background factors affect their language identities, and 3) to what extent does grade affect their language identities. The background theory consists of discussion of identity construction, language identity, the effects of CLIL and the connection between language identity and CLIL-education.

The data for this thesis was gathered via questionnaire. The sample consisted of 114 respondents from three IB-schools in Finland. The answers on the questionnaire items were quantitatively analyzed, and the answers on the open-ended question were analyzed qualitatively based on their content. The results showed that few students characterized themselves as monolinguals, whereas self-categorizations into groups of bilingual and multilingual were more frequent. Differences were found between all three groups in identification with language communities, feelings of bilingualism, others' views of oneself, and the effect of the number of languages used at home with other family members. Monolinguals did not seem to identify with any of the language communities offered in the questionnaire, whereas bilinguals tended to identify with bilingual speakers and multilinguals with multilingual speakers. Although no significant differences were established in the identity change of the students', the results showed the tendency of monolinguals having maintained their language identity, as non-monolinguals tended to feel like their language identity had changed during IB-studies. No differences between students on different grades were found. The participants' own descriptions of bilingualism were related to factors such as language competence, frequency of language use and affective factors. The results shed light on the subjective nature of language identity.

Further research could be conducted on the differences between monolinguals and non-monolinguals by gathering an even larger sample, and as identity is a changing construct, identity change of IB-students could be an interesting object of study. Also, monolinguals' identification with language communities was not established in this study, which could be studied in the future.

Key words: Bilingualism, CLIL, identity, language identity, language learning, social identity

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

Table of contents

1 Introduction.....	1
2 Identity	4
2.1 Social Identity and Language Identity	5
2.2 Monolingualism, Bilingualism, Multilingualism	8
2.3 Language Identity Categories: Descriptors of Bilingualism	13
3 CLIL (Content and Language Integrated Learning)	15
3.1 Effects of CLIL-teaching	16
3.2 IB-program as CLIL-teaching.....	18
4 Language Identity and CLIL.....	20
5 Data and Methods	25
5.1 Data gathering methods and participants	26
5.2 Questionnaire	27
5.3 Data Analysis	28
6 Results.....	29
6.1 Language identity of Finnish IB-students: monolinguals and non-monolinguals.....	30
6.2 Subjective experiences of bilingualism.....	36
6.2.1 Language competence, language use and others' views of one's bilingualism	36
6.2.2 Identity change	41
6.2.3 Subjective feelings of bilingualism	45
6.2.4 Identification with language communities	48
6.3 The effect of background factors	50
6.4 The differences in self-categorizations between grades.....	52
7 Conclusion	54
References.....	58
Appendix 1. Questionnaire for IB-students.	
Appendix 2. Finnish summary.	

List of figures

Figure 1. The structure of individual's identity.....	7
Figure 2. Overall distribution of self-categorizations.....	31
Figure 3. Participants with two mother tongues.....	32
Figure 4. The number of languages used at home with other family members (N).....	51
Figure 5. Self-categorization of the monolinguals and non-monolinguals on different grades(N).....	53

List of tables

Table 1. Definitions of bilingualism.....	14
Table 2. The distribution of language competence levels (%).....	33
Table 3. Subjective explanations of bilingualism related to the number of languages and age (N) by the participants.....	34
Table 4. Language learning and others' views of one's bilingualism/multilingualism...37	
Table 5. Subjective explanations of bilingualism related to language competence and language use (N) by the participants.....	38
Table 6. Confidence, identity and identity change during IB-studies.....	41
Table 7. Subjective explanations of bilingualism related to identity and code-switching (N) by the participants.....	43
Table 8. Subjective feelings of bilingualism.....	45
Table 9. Identification with different language communities.....	48
Table 10. The relationship between self-categorization and background (p).....	50
Table 11. Self-categorizations of bilinguals and multilinguals on different grades.....	54

List of abbreviations

CLIL	Content and Language Integrated Learning
IB	International Baccalaureate
IBO	International Baccalaureate Organization
DP	Diploma Programme
SLA	Second Language Acquisition
L2	Second language
EFL	English as a foreign language
ESL	English as a second language

1 Introduction

Before, English has been a foreign language in a Finnish society. Today, the role of English has grown, and it has gained more and more importance in all areas of everyday life: the status of English is remarkably high in today's professional life, education and popular culture (Leppänen & Nikula 2009). Furthermore, it has been predicted already at the beginning of 21st century that in the future, English is going to replace Swedish as the second language in Finland (Taavitsainen and Pahta 2003, 1). English language proficiency is not a rare skill anymore, but it is recognized as a necessity in all areas of life, including the education of the young. In Finland, IB-education (International Baccalaureate) offers an English-speaking education as an alternative to national schooling system. That is, as IB-education uses foreign languages in teaching, IB utilizes a form of CLIL-method (Content and Language Integrated Learning).

IBO (International Baccalaureate Organization) provides four different programs that can be taught in authorized schools (ibo.org). From four programs, the one that is an alternative to the national upper secondary education is known as Diploma Programme (ibid). The first class of Finnish DP started in 1990 (syk.fi). That is, IB-program has already been a part of the Finnish schooling system for 30 years. This year, over a thousand students are studying in Diploma Programs in Finland (The Union of Upper Secondary School Students in Finland, 2020). Still, very few studies have been conducted on IB-students in Finland.

Outside Finland, there have been studies about the effects of CLIL-teaching on students' language (e.g. Pérez, Luisa and Lancaster 2017) and content learning (e.g. Ruiz de Zarobe 2016), motivation and other subjects related to the psychology of students (ibid). In Finland, there have been very few studies related to IB-education. Also, even fewer studies have been conducted related to psychological effects of IB-education in Finland: McCambridge (2007) and Kovanen (2011) studied IB-students' language identity qualitatively with a small sample of interviewees, and as such, they did not offer generalizable results.

In this study, the aim is to create a more generalized view of Finnish IB-students' language identities. Language identity is a part of individual's social identity, which is formed by the feelings of belonging into an ingroup as opposed to an outgroup (Tajfel and Turner 1979). According to Tajfel and Turner's Social Identity Theory (1979), the formation process of social identity consists of three phases, through which every individual goes. The formation process of language identity follows the same pattern, and it is affected by multiple factors, such as the context of language learning (Higgins 2011).

Language identity is related to identification with different language communities (Edwards 2008, 27), and if an individual identifies oneself with more than one community, (s)he may possess a bilingual or multilingual identity (Grosjean 2008). In this study, bilingualism means regular use of two (or more) languages by an individual, whereas monolingualism means the use of one language, and multilingualism, again, means the regular use of multiple languages (ibid). Language identity is an important notion in second language acquisition (SLA) research as it is viewed as a psychological factor in language learning (Norton 2000). As the language identity of Finnish IB-students has not been studied much, it is a subject worth studying.

In this thesis, the aim was to find out what the Finnish IB-students' language identities are like, and which factors affect the formation of their language identities. Three research questions were formulated for studying this subject:

- 1) What are the language identities of Finnish IB-students like?
- 2) To what extent do the language identities differ depending on students' background and mother tongue?
- 3) To what extent do the language identities differ between students from pre-IB to 2nd grade?

The goal was to establish the differences between students according to their self-categorization into three language identity groups: monolinguals, bilinguals and multilinguals. These three groups were used as the most important way of dividing the students and comparing them to each other. By comparing these three groups, the differences in the effects of background factors, grade, and subjective experiences of bilingualism on students' language identities were established. The data was gathered by using a questionnaire, which was formulated based on Kasurinen's (2019) study on the experiences of bilingualism of Finnish university students of English and the research methodology by Dörnyei and Taguchi's (2009). All three research questions were answered by using quantitative methods of analysis. In addition, qualitative method was used to answer the first question more thoroughly by analyzing the contents of the participants' descriptions on the open-ended question about their language identity.

In the first theory section (section 2), an overall view of identity construction is given, as well as language identity being discussed in more depth. In the second theory section (section 3), CLIL-method is discussed, and IB-education is presented as an education using CLIL-method. In addition, in section 4, the connection between language identity and CLIL-teaching is discussed, and the previous studies conducted in the subject are presented. Data

gathering methods, the questionnaire used, and the methods of analysis are presented in detail in section 5. The results-section (section 6), in which the main results are presented and discussed, is divided into four parts: in the first subsection (6.1), the main characteristics of the language identity groups are presented. In section 6.2, the differences in subjective experiences of language identities between identity groups are presented in a thematical order, after which in section 6.3 the effect of background factors on self-categorization is analyzed. In the last section (6.4), the effect of grade on self-categorization is established. The findings are discussed in relation to theoretical background and previous studies already in section 6, and the conclusion, limitations of the study and suggestions for further research are provided in section 7.

2 Identity

Identity is a psychological property of humans, and it has been widely studied since 20th century. Erikson (1956) has made groundbreaking work in the subject, and according to him, identity can be seen as a personal maturation process: identity develops from birth to late adulthood. Although identity research has started from developmental psychology, SLA research has also focused on identity: in SLA, identity is seen as a changing construct for which the connection between the self and social environment is crucial and it is created in that relationship (Kanno 2003). The difference between these two is the main idea of what identity is and how it affects the individual: from the psychological perspective, the goal is to go through the maturation process and create a coherent “self”, which is an important developmental task in adolescence, when in SLA identity is more an affective and a personal aspect affecting language learning (Erikson 1956; Norton 2000). Furthermore, according to various theories (Erikson 1956; Marcia 1966) identity negotiation process occurs in adolescence and early adulthood, which could mean that the participants of this study, IB-students from the age of 16 to 19, are still forming their identity. This could affect the results in a way that some participants are further on their negotiation process than others and have created a more stable identity already.

Identity consists of *personal identity* and *social identity*. These two are intertwined in a complex manner, but most of the theories explaining identity structure define identity consisting of two main mechanisms: “individual or “self” identity (=personal identity, which includes role identity) and collective or “we” identity (=social identity, which includes group identities and social categories)” (Korostelina 2007, 35). Individual identity describes what makes people who they are and what makes them unique and different from others, and our habits and behaviors reflect the self (Saastamoinen 2006, 172). Social identity, as mentioned above, includes group identities and social categories, such as linguistic and cultural group memberships to which an individual belongs (Edwards 2008, 27).

Tajfel and Turner (1979) have also studied the continuum between individual and social identity, and they have defined social identity as the relationships between individuals that are affected by the individuals’ memberships in different groups and categories, and not affected by the relationships between different individuals. According to Thoits and Virshup (1997), individual identity means identifying oneself with a particular type of person, as social identity means that one identifies with a particular group of people that are similar based on the group members’ beliefs, interests and values. The development of one’s identity

is often seen as a process occurring in between these two poles (Korostelina 2007, 35). In some theoretical works, identity is divided into subcategories, as Korostelina (2007, 61) has done: identity is divided into core identities, short-term identities and situational identities:

“Core identities are fairly stable and dominant: they exist for a relatively long time and change only in situations of considerable social shifting; some core identities remain through an individual’s entire lifetime. Short-term identities are inconstant and reflect temporary ingroup and intergroup relations. Situational identities are connected with, and depend on, concrete situations. They are a “building material” for short-term and core identities” (Korostelina 2007, 61).

In other words, Korostelina (2007) suggests that identity formation consists of an unstable system of social identities, and different social identities affect one another. This is how individual’s asymmetrical identity system, which is open to change, is created. These three subcategories are related to constancy and can be considered as a road to a rather stable identity.

According to most theories, language identity is a type of social identity: it means that a person identifies themselves with a particular language group. Furthermore, language identity is created through the same processes as any identity. Since the focus of this study is on language identity, which is a subcategory of social identity, discussing individual identity and its development further is not relevant with more focus being on social identity, and even more on language identity, which will be discussed next.

2.1 Social Identity and Language Identity

The role of language in people’s mind has been studied quite much, and today language is seen as a window to the individual’s internal state (Bucholtz and Hall 2009, 20). Furthermore, the relationship between language and identity is thought to be as follows: “-- human identity is both constructed and constrained by language” (Evans 2018, 19). As stated earlier, SLA approach views identity as an affective factor on language learning, and language identity is something that is created in relation to the individual’s self and language (Norton 2000).

Henry Tajfel, one of the most famous researchers in social psychology, has developed *Social Identity Theory* (1979) that attempts to describe the nature of social identity and how it is formed. The basic idea is that social identity is the person’s sense of themselves based on the group they belong to: it can be for example family, volleyball team or school class, or in the case of this study, a language group. The formation of social identity occurs in an individual’s environment and in relation to that environment (Edwards 2009, 27). According

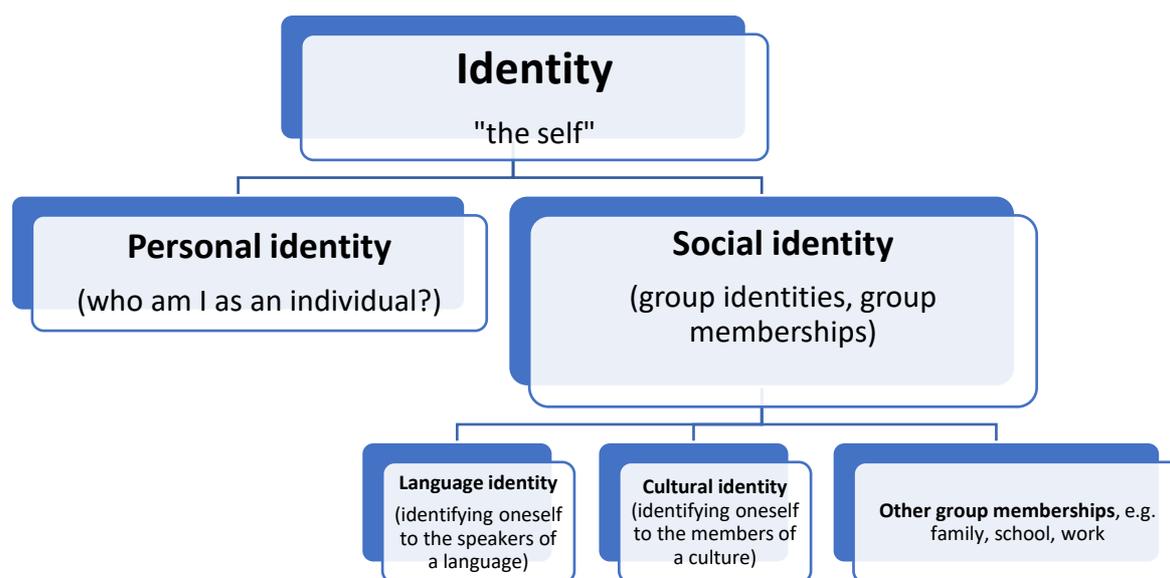
to Tajfel and Turner (1979), social identity is built upon individual's membership in an ingroup and the differences or comparison with members in an outgroup. This means that people identify themselves belonging to a specific group of people that are similar to oneself, and not belonging to other groups because one feels that (s)he is fundamentally different from the members of that group. This is a process called stereotyping, and it is a natural cognitive process in which people often tend to exaggerate. Furthermore, there is a possibility for an individual to make decisions about their adopted identities, and also how we make our identity visible to others: "--there are identities we embrace, identities we resist and identities we are unsure of or uncertain about either shared in a collective way or kept to ourselves in an intra-psychic way" (Wearing 2011, ix).

The formation of social identity is a great part of Tajfel and Turner's (1979) Social Identity Theory, and it consists of three mental processes that every individual goes through: *categorization*, *identification* and *comparison*. First, categorization means that we categorize ourselves and others to be able to understand our social environment: we behave according to the group we belong to and the group the other person belongs to. Second, identification means identifying ourselves with the group we categorized ourselves into and creating an identity accordingly, e.g. high school student identity. Identity created during the process will be emotionally important and affect one's self-esteem. The last stage, comparison, means the process of comparison with other groups. On this stage, we compare our ingroup with outgroups, which can sometimes lead to competition or hostilities. As opposed to how social identities are created, Korostelina suggests that it is possible to abandon already formed and internalized identities: "If a new identity begins to perform the necessary functions, it can lead to the disappearance of old identities. If one of the identities stops fulfilling its functions, it gradually loses its significance and vanishes" (2007, 62). This means that when an individual forms a new identity, if it is more important for the individual's functioning, old identities can be abandoned if they are not needed anymore. Thus, an individual can adopt new identities through a mental process, but (s)he can also abandon irrelevant identities along lifetime.

Language identity and cultural identity are subcategories of social identity (see Figure 1), and the negotiation of language identity is related to political arrangements, power relations and language ideologies, but also, as stated earlier, people's perceptions of each other's identities, and language is itself the mean for that identity negotiation (Jenkins 2007, 198; Tajfel and Turner 1979). Cultural identity, again, is related to identifying oneself with a cultural group in which the members share the same norms, values etc. (Hamers and Blanc

2000, 201). Nevertheless, often group memberships (and that way language and cultural identity) are based on dichotomies, e.g. native-non-native speaker. Often for example language identity is not something that fits into either/or-type of categorization, and many researches have tried to solve that problem. Kanno (2003) has, for example, suggested a new concept that does not fall into a dichotomy, but offers a new state for an individual to create their identity: the term is called *hybrid identity*. It provides an opportunity to describe identities that are a mixture of different categories and provides this study one more term for describing the participants language identities: hybrid identity is often relevant in English as a foreign language -contexts, where language and culture are mixed and the line between one identity or another is not too clear.

Figure 1. The structure of individual's identity.



In this study, the participants may come from different backgrounds with different L1s and different languages they use at home with other family members, but all of the participants use the same language at school: English. That is why hybrid identity is an important notion for this study. That is, in this study, language identity can be defined as one being on either monolingual/bilingual/multilingual-scale or defined having a hybrid identity, if an individual feels that (s)he does not fall into any of the preceding categories. Next, I will concentrate on definitions of monolingualism, bilingualism and multilingualism.

2.2 Monolingualism, Bilingualism, Multilingualism

Bilingualism is a debated subject, since there are different views about what bilingualism actually is and who is and who is not bilingual. According to some definitions there are more properties that need to be fulfilled in order to be called bilingual (the rest are monolingual), but if the concepts are defined in the simplest way, -lingual refers to language speaker and mono-, bi- and multi- describe the number of languages. That is, in everyday language, monolingual speaks one, bilingual two and multilingual many languages. Still, further discussion about the definitions of these concepts is needed.

From a psychological point of view, bilingualism has been defined as follows: “Bilingualism (multilingualism) refers to the coexistence of more than one language system within an individual, as contrasted to monolingualism” (Hakuta 2010, 174). Grosjean, again, defines bilingualism as “--the regular use of two or more languages (or dialects)” and bilinguals as “--those people who use two or more languages (or dialects) in their everyday lives” (2008, 10). According to these definitions, *multilingualism* has been defined as an extension of bilingualism: bilingual knows two languages (thus prefix bi-) and multilingual knows at least three languages or more (thus prefix multi-). Accordingly, that is how these three concepts are used in this study: monolingual here is an individual who knows one language system, bilingual is an individual who knows two language systems and multilingual is someone who is familiar with at least three language systems. Still, I will mostly discuss the concept of bilingualism when discussing language identity and its definitions, since it can be used in contrast to monolingualism, and also used instead of multilingualism as an extension of bilingualism. That is, multilingualism is a similar phenomenon to bilingualism, as similar factors are affecting multilingualism and bilingualism.

Bilingualism has different societal characteristics which may influence the level of bilingualism the individual is experiencing. Hakuta (2010, 175) lists 5 characteristics: social status, compartmentalization, literacy, immigrant generations and other historical circumstances. According to Hakuta (ibid.), *social status* refers to equal or non-equal status of both of the languages of the bilingual, but at the same time it can be used to refer to a situation called diglossia, which means that in given society, there are “different functions and domains of use for each language” (Bhatia&Ritchie 2012, 6). Thus, languages in that society have unequal status, as e.g. French vs. English in some parts of Canada (French as a minority language, but still official language of the country). Furthermore,

compartmentalization, too, means that the two languages have different functions and domains in everyday life (Hakuta 2010, 175), but compartmentalization is mostly related to individual bilingualism, of which an example could be from a Finnish setting: Kurds, which is a rather large group of immigrants in Finland, speak Kurdish as their mother tongue, but since Finnish is one of the official languages of Finland, they rarely get service in that language. That often leads to a situation where they either need to learn Finnish or use English in multiple situations, e.g. in stores and at school. As they become bilingual, it is possible that the languages stay compartmentalized, and Kurdish is the language of informal conversation, and Finnish becomes the language of formal situations, for example.

As the third characteristic Hakuta (2010, 175) mentions *literacy*, which means the level of literacy in both or all of the languages. In practice it can be related to a variety of reasons why someone is or is not equally literate in both of their languages, but according to Bhatia & Ritchie (2012, 13) “[o]ne might, for example, develop a fluent conversational grasp of a language in a relatively informal way, and only later feel the need to add some formal literacy skills”, which could characterize individual perceptions of their own bilingualism. As the last two characteristics Hakuta (2010, 175) has listed immigrant generations and other historical circumstances, e.g. colonization. In this context, *immigrant generation* is related to a language shift from heritage language to bilingualism and towards dominance of the new language, from which an example is given by Bhatia & Ritchie (2012, 6): “--classic pattern for immigrants to the United States has been bilingualism (mother tongue and English) by the second generation and English monolingualism by the third.” This way bilingualism can be thought as a pathway from one monolingualism to another monolingualism.

Furthermore, Hakuta (2010, 176-177) lists characteristics that describe bilingualism at an individual level that are often mentioned in literature when referring to bilinguals. One of the most important characteristics is *age of acquisition*. The age when an individual is exposed to a second language is one factor explaining the level of bilingualism attained (ibid). Hakuta differentiates two types of bilingualism: simultaneous, which means that already from birth individual acquires two languages, and sequential, which means that an individual begins acquiring second language later in life, for example in late childhood (ibid). Also Yip & Matthews (2007, 25) mention the distinction between simultaneous and successive (which means the same as sequential) bilingualism: in simultaneous bilingualism, a child receives regular input in two languages before the age of three, as successive

bilingualism means that a child is not exposed to a second language until after 3 years of age.

Another defining characteristic of bilinguals is *language proficiency*. According to Hakuta, a high level of language proficiency in two languages has been seen as a characteristic of bilinguals (ibid). Furthermore, Hakuta (ibid.) suggests that areas of language proficiency that describe bilingualism are grammatical proficiency (includes grammatical factors and phonology of a given language) and literacy. Broader view of competence has been provided by Bhatia & Ritchie (2012, 12):

“Given the four basic skills (listening, speaking, reading, and writing), and given further subdivisions so as to take into account divergent possibilities under headings like vocabulary, grammar, and pronunciation, we soon arrive at a considerable number of important elements. They will all figure in the assessment of language competence, but it does not follow that strength in one means strength in another --”

In other words, language proficiency consists of a variety of skills, and an individual may be unequally competent in different skills, since they are not connected to one another in a straightforward manner: for example, an individual can be good at reading, but incompetent in speaking in English. Therefore, it is rather difficult to give one definition of sufficient competence and being called bilingual depends on which level of competence is required. Furthermore, fluency could be added as one factor in defining language proficiency, as it is often thought to be a property of bilinguals (Grosjean, Li and Bialystok 2013, 11). Fluency is often used to refer to oral competence of the speaker, that is, speaking, listening and other skills related to oral communication, such as speech rate (De Jong 2018, 237). Still, the older the definition of bilingualism, the more restrictive it is regarding sufficient competence.

The third characteristic Hakuta mentions is the *active vs. passive bilingualism*, that describes the extent to which an individual uses their two languages, but that is not related to proficiency (2010, 177). This means that the use of two languages, the extent and the combination of usage can vary in different speech communities: for example, when an English person marries a Finnish person, they are likely to speak English, if the English-speaking person does not know Finnish. This way the Finnish person, if (s)he is living in Finland, speaks Finnish and English in everyday life and is an active bilingual. If Finnish is totally left out of the language repertoire, for example, in the case of living abroad and not needing Finnish every day, the Finnish person could be called passive bilingual, since English is the everyday language and Finnish is not often used. Still, there are many different

combinations of using different languages between family members, which depends on the context and people's own preferences.

Grosjean (2008) has divided bilingualism into four categories defining how bilingual people are viewed. That is, there are four types of social identity categories in which language identity and cultural identity are connected (see section 2.1). These views are monolingual view of bilingualism, bilingual view of bilingualism, monocultural bilingualism and bicultural bilingualism. According to Chen (2015, 1) "--language is regarded as a carrier of culture and a marker of group identity – language and culture are interrelated." This way, language and culture cannot be fully separated, even though there are four parts in Grosjean's model.

Monolingual view of bilingualism means, according to Grosjean, that the bilingual has "--two separate and isolated language competencies; these competencies are (or should be) similar to those of the two corresponding monolinguals; therefore, the bilingual is (or should be) two monolinguals in one person" (2008, 10). There are also a few consequences that characterize this view of bilingualism: among them are views of bilinguals having full fluency of the two languages and balance between these languages, not taking different needs and social functions for language use into consideration, full isolation of the two languages in a bilingual individual, and surprisingly, the bilinguals' own opinion about not having adequate fluency in both languages (Grosjean 2008, 10-14). Still, Bhatia & Ritchie (2012) have discussed whether bilinguals actually have two separate language systems, and have come to a conclusion that linguistic alteration is common for bilinguals:

"--switching may occur for emphasis, because the *mot juste* is found more readily in one language than in another, or because of a complicated network of perceptions of the speech situation, topic and content, the linguistic skills of interlocutors, degrees of intimacy and formality, and so on" (Bhatia & Ritchie 2012, 16).

What they are saying is that there are multiple reasons for word- and language choices during conversation, some of which may be related to the bilingual individual and some to other interlocutors. Thus, as this switching between languages during conversation, 'code-switching' or 'translanguaging', is a property of bilinguals, it cannot be said that full isolation of languages is required for an individual to be considered bilingual. Bilinguals' view of their own language competence is an interest of this study, too, since the participants of this study come from different backgrounds, and not all of them are native English speakers. Consequently, I would hypothesize that many of the non-native English speakers identify

themselves as non-bilinguals, as they do not think they are fluent enough in English, which is the language of their curriculum.

Bilingual view of bilingualism means that an individual is not the summation of two monolinguals, but a totally different composition. Grosjean (2008) states that “the co-existence and constant interaction of the two languages in the bilingual has produced a different but complete language system”, which means that the bilingual is “a unique speaker-hearer” of the two languages (2008, 13-14). According to this view, bilingual is fully competent in two languages and the skills that the bilingual has reached are at the required level of competence in order to fulfill the needs of the individual and the environment (ibid). Bilinguals can also use the combination of the two languages or one of them in different situations, with different people or for different purposes (Grosjean 2008, 14; Bhatia & Ritchie 2012, 16). Grosjean continues by saying that it is acceptable for a bilingual to not be as competent in both of the languages that (s)he is speaking: “Because the needs and uses of the two languages are usually quite different, the bilingual is rarely equally or completely fluent in the two languages.” (2008, 15). Furthermore, being called bilingual does not require being as fluent as monolinguals in those languages: “If one were to count as bilingual only those who can pass as monolinguals in each language, one would have no label for the vast majority of people who use two or more languages regularly but do not have native-like fluency in each.” (Grosjean 2010, 20). In addition, native-like competence is a problematic concept in describing bilingualism as it is. To conclude, bilinguals do not have to be as fluent in both languages to be called bilingual, but often high proficiency, including fluency, is what is expected of bilinguals, as was mentioned earlier when discussing the characteristics of bilingualism (Grosjean, Li and Bialystok 2013; Hakuta 2010).

Biculturalism is often defined in the same manner as bilingualism in everyday language, but it is connected to identifying oneself to more cultures than just one (Grosjean 2008, 214). Grosjean (2008, 214) has defined biculturalism with 3 characteristics: biculturals “--take part, to varying degrees, in the life of two or more cultures--adapt, at least in part, their attitudes, behaviors, values, languages, etc., to these cultures--combine and blend aspects of the cultures involved.” Chen (2015, 3) continues this by saying that “--just as bilinguals differ in the context in which they learn their two languages, bicultural individuals also vary in the context in which they acquire their second culture.” Furthermore, bilingualism and biculturalism are not necessarily comparable, since bilingual may not be bicultural and bicultural may not be bilingual (Grosjean 2008, 217). Still, often people who could be called

bilinguals in principle, feel like they do not identify themselves with the other culture, although language has great effect on expressing cultural norms and beliefs. (Grosjean 2008, 217; Luna, Ringberg, Peracchio 2010, 78). It is possible, then, to be called *monocultural bilingual*: they are “--those individuals who never internalized the culture attached to their second language. They typically learned their second language in a classroom environment, without significant exposure to the corresponding cultural context” (Luna, Ringberg, Peracchio 2010, 78). This is a concept in opposition to *bicultural bilinguals*, who are “those individuals who have internalized two cultures and who speak the languages associated with each of those cultures” (Luna, Ringberg, Peracchio 2008, 279). That is, biculturalism is mostly related to individual perceptions of biculturalism and bilingualism: “--what defines a bicultural individual may be extended from extensive experiences in two cultures to bicultural identifications coupled with bilingual competence” (Chen 2015, 4). Becoming bicultural is a process that involves being in touch with two or more cultures and, at least in part, having to live within those cultures (Grosjean 2008, 217). This can happen already in the early childhood, when a child is impacted by two or more cultures at the same time, and the impact can continue throughout life (ibid). Furthermore, it is possible to become bicultural later in life, because of e.g. migration (ibid). Thus, identifying oneself with the other culture is related to individuals’ experiences about being bilingual, although from the Grosjean’s point of view, bilinguals can be either monocultural or bicultural, and still be called bilinguals.

In this study, bilingualism is defined as the existence of two language systems within an individual, because that defines bilingualism clearly in a broader sense but leaves room to different definitions of bilingualism, such as age of acquisition and level of proficiency. As bilingualism is a part of one’s identity, participants’ subjective definitions of bilingualism are taken into consideration when analyzing whether someone is or is not a bilingual. Furthermore, a broader and more permissive definition of bilingualism is more applicable in a study that discusses identity, because as identity is a subjective perception of oneself as a person, it is not relevant to draw a strict line between monolingualism and bilingualism.

2.3 Language Identity Categories: Descriptors of Bilingualism

In this study, there are three main categories of language identity: monolingual, bilingual and multilingual. According to theoretical works presented in the previous chapters, I have conducted a categorization for language identity based on definitions of those three concepts. In this categorization, descriptors of bilingualism such as the age of acquisition (early or

late), fluency in the two languages or the level of competence, affective factors such as emotions towards the language community or other social factors e.g. immigration and the frequency of the L2 usage are taken into consideration. As this thesis studies language identity, it is important to include identification as one definition of bilingualism: identifying oneself with L2 community for any reason or defining oneself bicultural bilingual or monocultural bilingual. Below is a table in which all the definitions have been gathered.

Table 1. Definitions of bilingualism.

Factor of bilingualism	Descriptor
Age	Simultaneous vs. Sequential bilingualism (Early vs. late bilingualism), age of acquisition as a descriptor of bilingualism
Proficiency	Nativeness/non-nativeness, fluency and the levels of competence of the languages as descriptors of bilingualism
L2 usage	Frequency of language use and compartmentalization of languages as descriptors of bilingualism
Affective factors	Emotions towards language communities as a descriptor of bilingualism, other social factors (e.g. immigration) as descriptors of bilingualism
Identification	Identifying oneself to language communities and cultures as descriptors of bilingualism
The number of languages	More than one language at satisfactory level of competence as a descriptor of bilingualism

Some students may have English as their first language, or they have learned English in different ages and at a different pace, which could affect their language identity (e.g. Hakuta 2010) Also, some may feel near-native and that way bilingual, and other may feel that they are not fluent enough in English to be called bilinguals(e.g. Grosjean 2008; Grosjean, Li and Bialystok 2013; Bhatia & Ritchie 2012). Emotional explanations can also result in defining oneself as bilingual. What comes to identification, students may identify themselves with English language community because of a variety of reasons, but also because of cultural

ones: identifying themselves with both Finnish and English cultures could result in bilingualism (e.g. Grosjean 2008; Chen 2015; Luna, Ringberg, Peracchio 2010).

There is a variety of descriptors of bilingualism in defining whether one falls into one category or another, and these descriptors have been presented above in Table 1. In this study, bilingualism is used in an opposition to monolingualism: if no requirements of bilingualism are fulfilled, the individual most likely is monolingual. An extension of bilingualism, then, is multilingualism, where the requirements of bilingualism are fulfilled in more than two languages. To be able to categorize individuals who do not feel that they belong to any of the three categories alone, hybrid identity is the fourth possible identity category (Kanno 2003). Therefore, participants can be categorized into four groups: monolingual, bilingual, multilingual or hybrid identity. Based on these categories and definitions, I will find out how the participants define themselves in relation to their linguistic identity and why.

Now that the definitions of bilingualism and language identity have been given, one possibly influential factor in creating bilingual identities should be discussed next. In the next section, I will present the notion of CLIL, discuss whether it could have an effect on language identity because of its nature, and present IB-program as a type of CLIL.

3 CLIL (Content and Language Integrated Learning)

CLIL is an approach in education that has received a lot of interest during the last few decades especially in Europe. CLIL, as is mentioned in the heading, comes from Content and Language Integrated Learning. In the Finnish education system, children are being taught in their mother tongue, which is usually either Finnish or Swedish, if immigrants are not taken into account. In CLIL, the goal is to “--integrate language learning with content learning, usually by careful coordination of both types of input, or by focusing on the acquisition of skills needed to cope with both areas” (Breeze et al. 2014, vii). In other words and in relation to the Finnish context, it means that teaching is done in a foreign language, which in this study is English. Furthermore, the role of foreign language is explained as being the center of this mode of teaching, in which foreign language is used for all kinds of communication, learning and interaction in the classroom (Halbach 2014, 1). Still, according to Nikula (2007, 208-209), CLIL is not a specific teaching method, but is used as an umbrella term that describes different types of instruction that is mediated through foreign language. In this study, CLIL is viewed as an approach of using a foreign language, here English, as the medium of instruction, which is used in IB-education across the country.

The effects of using a foreign language as a means for communication in classrooms have been studied widely for example in Spain, where it has been noticed that learning English to a satisfactory level is hard and almost never achieved, although children start learning language already in childhood (Halbach 2014, 2). Halbach (ibid.) describes the effect of CLIL teaching as follows:

“This focus on the language as an object of study changes radically in a CLIL context where the foreign language becomes a working tool and a means for communication about other contents. Students need the language to be able to follow their content lessons and participate in them successfully, and this in turn means that their approach to, and motivation for, studying the foreign language changes drastically.”

In other words, in CLIL contexts the motivation for learning a foreign language is driven by “the desire to survive” in the classroom environment. It is important to achieve a satisfactory level of language competence in that specific language to be able to learn the contents that are being taught, but also to be able to follow and participate in interaction in the classroom and often even outside classroom with teachers and other students. As the focus is rather on language use than on language itself, in theory it would be possible to achieve quite high communicative competence in this kind of learning environment.

As has happened all around the world, globalization, international trading and demographic changes, such as immigration, have influenced the linguistic changes in Finland, too. According to Taavitsainen and Pahta (2003, 1), it has been a trend already early in the 21st century that Nordic countries are slowly changing from ‘English as a foreign language’ or ‘EFL’ societies toward ‘English as a second language’ or ‘ESL’ societies. This has been an ongoing process, and today English plays a crucial role in business, professional life, education and popular media (Leppänen & Nikula 2009). Also, according to Pyykkö (2017), good proficiency in English has become an important skill in working life. Thus, it could be that the status of English in Finland, which has been growing, has affected the rise in interest in CLIL-teaching in Finland. In the following sections, I will discuss the effects of CLIL that have been found by researching CLIL environments and present IB-education as a type of CLIL-method.

3.1 Effects of CLIL-teaching

There have been many studies about the impacts of CLIL on students’ language skills, content learning and teachers’ beliefs, but not that many of the impacts on language identity, which is the basis of this study. As the foreign language is the center of all communication

and learning, it is safe to say that it will have at least some effects on students as they create a whole different world where another language is used.

As noted earlier, using L2 in all learning situations could result in higher motivation for learning the language in order to learn the contents, but it has also been proved that it could result in using the L2 as a starting point of all communication and learning: there is no longer the need for using mother tongue, as students in bilingual degree in teacher training in Spain have told (Halbach 2014, 2). Furthermore, Coyle (2007, 553) suggested that in CLIL environments students have to learn new communication and learning strategies, as the language used by the teachers is not their mother tongue and the same strategies do not work – for example, how to guess the meaning of foreign words.

Pérez, Luisa and Lancaster (2017) studied the effects of CLIL on two areas of language skills: oral comprehension and production in L2. There were 24 participants, who were in the fourth grade of compulsory secondary school in Spain, and from which half were in CLIL-class and half received regular teaching. In opposition to a cross-sectional study on the effects of CLIL, they wanted to conduct a longitudinal study to reveal the development on oral English language skills of the participants in CLIL and their non-CLIL counterparts over the course of 1,5 years. What they found out was that in the test, CLIL students showed “--more developed oral receptive skills on the overall mark of the test and specifically for the most cognitively demanding task, thereby corroborating prior research claiming CLIL to enhance the understanding of complex language in a listening context” (2017, 316). After the period of longitudinal study, CLIL students showed superiority in fluency, grammatical and lexical skills (ibid).

Pérez and Luisa also studied the effects on the L1 and content learning in monolingual contexts. They had over 2 000 participants from both primary and secondary education, and found out that CLIL does not have negative effects on children’s L1 competence, but it does not degrade content learning either (Pérez and Luisa 2018, 1): actually, as has been studied by Surmont et al. (2016, 319), CLIL students may even outperform their traditionally schooled peers in different school subjects, such as mathematics. What has been found out by Surmont et al. (2016) and Pérez and Luisa (2018) is that using L2 in learning situations improves metalinguistic awareness and the improvement can extend beyond linguistic domain. This suggests that CLIL would have positive effects on students’ language abilities but also on content learning.

Juan-Garau and Salazar-Noguera (2016) discuss the different aspects of CLIL, the effects of CLIL and present different kinds of research that have been conducted on the area of

study. Ruiz de Zarobe (2016, 51-62) presented different studies that have tried to find out whether CLIL-students actually learn contents better than their non-CLIL counterparts, and it has resulted in controversial results: some studies show outperformance of CLIL-students, some studies show no crucial differences and some studies even have resulted in negative results when learning contents in a foreign language. To conclude, there are different views of how CLIL actually affects students and their learning, but mostly the results that researchers have got lean towards more positive results than negative. Among them are e.g. increase in confidence, higher motivation, feeling special as a learner, and that it challenges learners in a different way than traditional education system, without harming content learning (Juan-Garau and Salazar-Noguera 2016, 62).

Furthermore, according to Vollmer et al. in Juan-Garau and Salazar-Noguera (2016, 55), CLIL-students could reach a higher level of tolerance of frustration, which can be helpful in learning situations. Here it is important to remember that individual differences are a crucial factor when it comes to how CLIL affects students: some students may be more determined to learn the language or contents, which could result in positive outcomes in these studies. As opposed to positive results, individual differences may be behind negative results, too: low confidence, low motivation, and getting frustrated easily may result in negative outcomes. Moreover, there are many external and circumstantial factors that can result in a success or a failure of CLIL-teaching: among them are, for example, educational standards in a specific country, school and subject.

As I am trying to find out how Finnish IB-students identify themselves linguistically, it should be clear that CLIL-environment may have some effects on their identities as they live in an environment where they have to use another language that, for most of them, is a foreign language. In addition to effects on learning and motivation, challenging learning environment may affect language identity, which is a part of social identity of an individual. As has been addressed earlier, language identity of a bilingual can be defined by e.g. native-like fluency in two languages, which can be the case for some IB-students, thanks to CLIL-method. Next I will explain the fundamentals of the IB-program in Finland and why it can be considered as an education that uses CLIL-methods.

3.2 IB-program as CLIL-teaching

IB-organization (International Baccalaureate Organization, IBO) is a worldwide organization that provides education for children aged 3 to 19. In addition to the IB Diploma Programme, that is similar to the Finnish upper secondary school and to which in this thesis

I have referred as IB-education/program/school, they offer Primary Years Programme (3-12 years), Middle Years Programme (11-16 years) and Career-Related Programme (16-19 years, similar to vocational school) (ibo.org). In Finland, there are 16 schools that provide IB-education (syk.fi).

IB Diploma Programme, which is the object of this study, is a program offered in Finland for a second-degree education beside national upper secondary school. Although Diploma Programme (DP) itself is a two-year education, in Finland the students are provided with pre-IB grade, that is similar to 1st grade of national upper secondary school. After pre-IB, students can still choose to go the national upper secondary school or continue with DP. All schools providing IB-programs have to be authorized and follow a specific curriculum, which is carried out in a foreign language. This means that all subjects belonging to the curriculum are taught in a foreign language, therefore using CLIL-method. At the moment, IB-programs are offered in three languages: English, French and Spanish, from which English is the language that is used in Finnish DP.

In short, the curriculum consists of three core elements and subjects from six subject groups. The three core elements are compulsory to all: theory of knowledge (reflecting the nature of knowledge), extended essay (research paper) and creativity, activity and service (project). In addition, all students have to choose subjects from the following six subject groups: studies in language and literature, language acquisition, individuals and societies, sciences, mathematics, the arts (ibo.org). Subjects are taught at two levels - high and standard level – and students have to choose three subjects at high level, the rest can be standard level courses (ibo.org). According to IBO's website, IB's programs differ from national curricula because they:

“encourage students of all ages to think critically and challenge assumptions [;] develop independently of government and national systems, incorporating quality practice from research and our global community of schools [;] encourage students of all ages to consider both local and global contexts [;] develop multilingual students” (ibo.org).

Especially the last one, developing multilingual students, is a goal that is related to the goals of this study: as IB's programs aim at developing multilingual students, it is important to find out whether it is actually happening. Still, as language identity is a subjective perception of oneself as a language speaker, students claiming to be monolingual cannot be viewed as a failure in creating multilingual students. More important is to find out how students explain their language identities and whether there are any remarks to be made about how Finnish IB-students view themselves as language speakers and why. Furthermore, IB's program's

philosophy includes basic values, that are summed up in their “What is an IB education” text written for everyone interested in IB education:

“The International Baccalaureate aims to develop inquiring, knowledgeable and caring young people who help to create a better and more peaceful world through intercultural understanding and respect. -- These programmes encourage students across the world to become active, compassionate and lifelong learners who understand that other people, with their differences, can also be right” (ibo.org).

Therefore, IB education is focused on upbringing stable young people and promoting healthy identity formation through challenging education.

In addition to the goals of IB’s programs, the reason for creating international education system and IBO’s philosophy is as follows:

“ --the Diploma Programme -- sought to provide a challenging yet balanced education that would facilitate geographic and cultural mobility by providing an internationally recognized university entrance qualification that would also serve the deeper purpose of promoting intercultural understanding and respect” (ibo.org).

This means that IB-programs try to bring people from different nationalities closer to each other and help for example immigrants in accommodating and students in gaining eligibility for further studies in different countries. As these Diploma Programs (DP’s) try to educate students to critical, multilingual thinkers, it should be noted that the program probably has an effect on students’ identity formation, which is still happening at the age of 16 to 19, when students go to DP.

To conclude, IB Diploma Program is a program which uses CLIL-teaching methods, and which promotes healthy identity formation and multilingualism. As DP is fundamentally different from national schooling system, it could be argued that the effects on students are different in these two systems, too. Furthermore, as in DP, many descriptors of bilingualism (see Table 1) could be found, it is interesting to see whether students who have not grown up using English actually feel that they could be called bilinguals. When it comes to language identity of IB-students, there have been few and rather small-scale studies on the subject, which all have had a discursive and/or comparative point of view in analyzing identity. In this study, interest is directed at a larger number of IB-students than in previous studies and the aim is to create a representation of language identity of IB-students without comparing IB-students to any other group.

4 Language Identity and CLIL

In earlier sections I have discussed the definitions of language identity, how it is a part of social identity and how it is created. I have also presented the definition of CLIL, its effects

on learners and how Finnish IB-program is an example of CLIL-teaching. As I am trying to find out whether Finnish IB-student's language identity is affected by CLIL-environment, I have to form a clear connection between language identity and CLIL by discussing how CLIL may change students' subjective opinions on their own identity and present some of the previous studies on the subject.

As I have presented earlier in section 2, identity is a construct that is negotiated and created in the relationship of the self and social environment, and that communication between the self and social environment and relations is done through language (Kanno 2003). This means that as language is a tool for identity negotiation and it gives the learner power to choose and modify their identity, learning a foreign language gives more power for the learner to choose their language identity (Kanno 2003, 14). As language identity is a part of the individual's social identity, it reflects the individual's subjective feeling of belonging to a language community and/or culture (e.g. Kanno 2003; Tajfel and Turner 1979; Grosjean 2008). Thus, language identity is a subjective perception of belonging to a language community, and it is created through negotiation between the self and social environment, which is done by using language. That way, learning new languages can affect an individual by changing one's language identity based on this new power that the individual has gained. Learning additional languages well enough can enable negotiating bilingual or multilingual identities (e.g. Hakuta 2010; Grosjean 2008).

One more factor affecting language identity negotiation is the social and cultural context of the language learning process (Higgins 2011). That means that language learning processes are connected to the language identity negotiation by their physical context, which in this case, is school. As in CLIL a foreign language is used as the language of education, it gives opportunities to communicate in that language. Furthermore, individuals who treat bilingualism as a synonym to native-like competence, may form their language identity based on only that view of bilingualism. That is, native-like competence in common language is used as a synonym to the ultimate level of proficiency in a given language. According to Pavlenko (2003, 257-259), for example, native-like competence as the ultimate goal may cause feelings of incapability when achieving the membership of the native speaker community and if it cannot be fully achieved, it may cause feelings of not being a legitimate member of the speaker community. Even though bilingualism is not described on native/non-native-scale and 'bilingual' does not mean a native speaker of two (or more) languages, students who aim at native-like competence and regard that as the most important descriptor

of bilingualism may categorize themselves as monolinguals if the sufficient level of proficiency is not achieved.

Furthermore, language identity negotiation is related to a notion of investment: according to Norton Peirce (1995, 18), using the target language is not only for exchanging thoughts with native speakers, but also for constant organizing of the self and their relation to their social environment. That way, learner is investing in their own social identity at the same time when investing in the target language (ibid.) To conclude, language identity negotiation is related to the context of learning and investment in target language, both of which can be found at CLIL-school. That way, attending to a CLIL-program could be a crucial factor in identity formation and change.

CLIL-method, as discussed earlier in this section and in section 3, is a teaching method, where teaching is carried out in a foreign language (Breeze 2014). This way the whole education system forms an environment which promotes foreign language learning, and which focuses on language use and learning contents in another language than the student's mother tongue (ibid.) The effects of CLIL-method have been proved to be positive and mostly resulting in a good content learning and high competence in the language used at school (Breeze 2014; Juan-Garau and Salazar-Noguera 2016). As students usually learn the foreign language little by little as their education proceeds, by the time of graduation it could be that students feel that their language identity has changed thanks to a higher proficiency in English. Since fluency and affective factors are two possible defining factors of bilingualism among others (Hakuta 2010), after 3 years of education received in English may affect one's language identity by changing it from feeling monolingual to bilingual. Still, it is important to remember that some students studying in IB may already have English as their first language, and that way they may feel that their language identity has not changed during their studies. Thus, it is important to take different backgrounds into consideration: this study is conducted to find out what the identities of students in Finnish IB-schools are like in general, not just how IB-students with Finnish as a mother tongue describe their language identities. That is what makes this study different from other studies done on CLIL-method and language identity in Finland.

IB-education has been gaining more and more interest among the young, as it challenges students in a different way and opens doors for later studies. Still, there are very few studies related to IB-education in Finland, and especially to the psychological aspects regarding IB-education. A few studies have been conducted on CLIL and its impacts.

Roiha (2019) has recently studied former pupils that had received CLIL-education and concentrated on effects on their English language learning and self-concept, attitudes and life courses, which is close to the subject of this study, as language identity is created in relation to these factors (see section 2.2 and 2.3). What he suggests is that CLIL may have a crucial role in creating positive attitudes towards the English language and in forming the English language self-concept. Roiha's (2019) study was a qualitative case study (n=26) and the focus was on former pupils' experiences of the influences on self-concept, which includes general self-perceptions and domain-specific self-perceptions of an individual: that is, self-concept is mostly related to such perceptions and self-beliefs as self-efficacy (2019, 34-40). Thus, self-concept is a partly similar concept to identity ('the self', as described in section 2), but ultimately is substantially different as it does not take identification into consideration. Therefore, it is reasonable to move on to studying language identity and the identification of students receiving CLIL-education on a larger scale.

CLIL-education in upper secondary level is challenging in a different way and the education highlights the importance of research, it is more similar to university studies than regular high school. There have been studies on language identity in Finland, for example MA thesis on the experiences of bilingualism of Finnish university students (Kasurinen 2019), but very few studies have concentrated only on Finnish IB-students' language identity. Moreover, studies on the effects of CLIL-teaching on pupils self-concept (e.g. Roiha 2019) and teachers' identity negotiation (e.g. Pappa 2018) have been conducted in Finland, too, but very few have concentrated on language identity and more specifically bilingualism perceived by students that have received CLIL-education on DP level.

McCambridge (2007) studied the identities of graduates from English School of Helsinki. She used semi-structured interview (n=7) and analyzed the interviews using discourse analysis method. The main results were that English played a role in participants' identities, but they did not identify themselves with native-speaker communities, although they felt more multicultural than other Finnish people. All of the participants had Finnish parents and had studied in English for most of their lives, which would suggest that these two factors would have affected the similar results between participants. In this study, the sample is larger, and the participants very likely come from different backgrounds, so that it could be possible to make generalizable assumptions of the effects of different backgrounds on students' language identities.

Another study conducted was Kovanen's (2011) content analysis based on semi-structured group interview of Finnish IB-students (n=6). She wanted to find out how IB-

students viewed themselves as English language learners, and that way analyzed their language learner identities. She found out that although English was a big part of the participants' everyday lives, they did not identify themselves as multilinguals. As Kovanen focused on language use and the importance of English and because language identity was felt as a difficult subject for teenagers to describe, language identity was a rather small part of the study.

These two studies touch my study in a way that both McCambridge (2007) and Kovanen (2011) studied language identity of IB-students to some extent, and both have found out that in opposition to students in national schools IB-students feel more international and fluent in English, but still do not feel multilingual. As any studies have not been conducted on a larger scale in multiple Finnish IB-schools, the results of the previous studies cannot be generalized to describe the language identities of IB-students in Finland. That is why I wanted to find out how IB-education affects the language identity of the young and how they explain their feelings of belonging to a language community of their choice. In addition, as most of the studies on IB-students and CLIL-education have been based on interviews and have used discourse analysis and/or comparative measures in analyzing students' experiences in a qualitative manner, I decided to study a larger number of students and use both quantitative and qualitative methods.

In this study, I want to find out whether IB-students studying in Finnish IB-schools identify themselves monolingual, bilingual or even multilingual, and try to find out why with the help of the categorizations presented in section 2.3, but also try to find out whether their schooling system, in which instruction is fully in English, has had an effect on their identity, since it is connected at least to the frequency of L2 usage and compartmentalization of languages and that way could result in subjective feelings of bilingualism. The research questions are as follows:

- 1) What are the language identities of Finnish IB-students like?
- 2) To what extent do the language identities differ depending on students' backgrounds and mother tongue?
- 3) To what extent do the language identities differ between students from pre-IB to 2nd grade?

The first research question is rather clear in a way that with the help of the questionnaire, I am trying to find out whether students feel monolingual, bilingual or multilingual and why,

or why not. As identity is a subjective construction that is formed and reformed throughout lifespan, it is important to ask students' own opinions on the matter. I would hypothesize that most of the participants categorize themselves as monolinguals, since they do not possess native-like competence in English. It could be possible, though, that some part of the students categorizes themselves as bilingual or multilingual, as they have attained a satisfactory level of competence on at least English, thanks to schooling system that provides English language teaching from primary school and IB-education.

The second research question is for finding out whether background has any effects on language identity: how do students who have gone to Finnish primary and secondary school and students who have immigrated to Finland at some point of their childhood categorize themselves? What about mother tongue, are there any consistencies in participants' answers based on their mother tongue? I would hypothesize that students who have gone to school in English categorize themselves as bilingual or multilingual, and students who do not have any connections to English-speaking communities would categorize themselves as monolinguals. There are other factors, such as mother tongue of the parents and studying abroad that could affect the participant's language identity, too.

Lastly, with the third research question I am trying to shed light on language identity change during IB-education. As this is not a longitudinal study but a cross-sectional one for time-related reasons, no clear assumptions of changes on language identity because of school can be made. Still, it would be interesting to see whether there are some convergences on participants' answers according to the grade at which they are at the point of answering the questionnaire. I would hypothesize that students on 3rd grade of DP are more likely to answer being bilingual or multilingual than students on the 1st grade (pre-IB).

As I have now discussed IB-school's possible connection with language identity of the young through IB education's basic philosophy and teaching methods, as well as presented the research questions of this study, in the next section I will discuss data gathering methods that I have chosen for this study.

5 Data and Methods

In the present study, I find out what the language identities of IB-students in Finnish IB-schools are like and to what extent do they differ in regards of students' background and years of study. By choosing both quantitative and qualitative methods, I establish generalizable information about how IB-students view themselves as language users. First, I will discuss data gathering methods chosen for this study. Next, I will present the

questionnaire and explain how it was conducted. In the last section, I will describe how data analysis was done.

5.1 Data gathering methods and participants

As language identity has mostly been studied qualitatively through interviews, any generalized knowledge of today's IB-students has not been established. Thus, quantitative data from a larger group of students was needed, and as a data gathering method, I used a questionnaire that includes multiple choice and Likert-scale questions: that way I investigate language identity and students' experiences of bilingualism and make assumptions based on the answers of larger student populations (Dörnyei and Csizér 2012, 74-75). Furthermore, as questionnaires are also a way to establish the effects of background factors, it was a suitable method for data gathering in this study, too (*ibid.*) Still, as identity is a complex, subjective phenomenon and students may feel differently of what makes their language identity, I wanted to give students an opportunity to describe themselves what their language identity is like. Therefore, I included one open-ended question at the end of the questionnaire to avoid simplifying and objectifying this very subjective and complex phenomenon too much (Dörnyei and Taguchi 2009, 7).

To be able to get valid data, I comprised a questionnaire with questions as simple as possible and also provide short definitions of the most important concepts, mono-, bi- and multilingualism, to be able to avoid the situation where participants do not understand the concepts and that way do not know how to answer the questionnaire truthfully. Furthermore, I comprised a questionnaire which is short enough for students to answer without getting tired, but which is long enough to gather enough data from the participants opinions (Dörnyei and Taguchi 2009, 8-9). As questionnaires are commonly used in SLA and psychology research and they are suitable for gathering large amount of quantitative data that does not aim at going into detail of language identity construction but aims at creating generalizable data, questionnaire was the most well-founded data gathering method for my study (Dörnyei and Csizér 2012, 74). The questionnaire was piloted by eight respondents from different Finnish universities and polytechnics.

The link to the Webropol questionnaire was sent to a few IB-schools in Finland via email, which means that the study is based on cluster sampling: one IB-school represents a subgroup of the whole IB-student population in Finland. Still, students had the opportunity to opt out, since answering the questionnaire was fully voluntary. Every school either filled up the questionnaire during a lesson, or the link was shared with students online. The link

was shared with all IB-students from pre-IB to 2nd grade of Diploma Programme. Since the sample is formed of three different schools in Finland, and as there are approximately 1000 IB-students in Finland (The Union of Upper Secondary School Students in Finland, 2020), the participants (N= 114) could be viewed as representative of the whole group of students in Finland. That way, using a questionnaire allowed me to find some generalizations of IB-students language identity. Students in IB are 16 to 19 years old, and they are likely to come from different backgrounds. Some have lived abroad as some have started studying in English at the upper secondary level. Furthermore, it is likely that students use different languages in different situations, such as with family members. Thus, it is important to take background information into consideration.

As language identity is an abstract concept, I operationalized it with Likert-scale question patterns that were related to different descriptions of bilingualism (see Table 1). The data was gathered during autumn 2020, and it included only answering the questionnaire fully anonymously online without any follow-ups. In the next section, I will discuss the questionnaire in more depth.

5.2 Questionnaire

The questionnaire was conducted based on Kasurinen's (2019) questionnaire of subjective experiences of bilingualism of Finnish university students, but it was modified to apply to my research questions with the help of Dörnyei and Taguchi's (2009) work. It included overall 16 questions or question patterns, from which 1 was an open-ended question where students were supposed to continue the sentence "In my opinion, I am a bilingual because..." or "In my opinion, I am not a bilingual because...". To avoid the issue of leaving the open-ended question blank, it was marked as compulsory, but the problem of giving too brief answers could not be controlled beforehand (Dörnyei and Taguchi 2009, 10). As language identity is a subjective view of oneself as a language user and a member of language community, I added an open-ended question to provide participants with an opportunity to describe their subjective experiences in their own words, to say something about their language identities that the questionnaire did not cover and to see, which is the most important descriptor of bilingualism in their opinion.

From the remaining 15 questions, 10 questions were multiple choice questions which were designed to gather background information and 4 were Likert-scale question patterns with similar themes relating to language identity in each pattern. Background information consisted of multiple-choice questions of grade, mother tongue, strongest language, second

strongest language, third strongest language, languages used at home with family members, the age at which the participants started learning English, whether the participant attended English speaking schools and whether the participant has stayed abroad and for how long. These questions were chosen to establish participants' backgrounds and possibly to see whether different backgrounds have any visible effects on participants' language identities.

On the second page, there was a brief definition of bilingualism to make sure that all participants have some knowledge of what bilingualism actually means. After that, there were four Likert-scale question patterns with overall 26 items related to language, language use and language identity. Items were constructed to in a consistent manner: all were affirmative statements, started with similar pattern such as "In my opinion..." and "I feel...". The items in the first pattern were related to English language use, the second included items of identity as a language speaker, in the third section items established feelings of bilingualism, and in the last set the questions were designed to find out to which language communities participants identify themselves. Likert-scale questions were on four-point scale: strongly disagree, disagree, agree and strongly agree.

Furthermore, one multiple choice question was added to directly ask with no explanations needed, whether the participant identifies themselves as mono-, bi- or multilingual. The question was placed at the end of the questionnaire, so that participants would be prone to answer the question regarding their language identity in a straightforward manner after answering questions that are related to the construction of that identity.

The reason for using three different types of questions was to obtain results that describe the different sides of language identity. Answering the questionnaire took approximately 5 to 10 minutes, and as the participants are studying in an international program and come from different backgrounds, the questionnaire was in English. The whole questionnaire can be found from the appendix 1. The quantitative analysis and the data received from the questionnaire were used to find what the language identities of IB-students are like and to what extent are there similarities between students from different backgrounds and between pre-IB students and 2nd graders.

5.3 Data Analysis

Because of the nature of this study and mostly quantitative data, data analysis was done statistically by using IBM SPSS Statistics -program. First, I organized the quantitative data in Excel, after which I analyzed it using SPSS. The sample was divided into three groups according to the participants' self-categorization – whether they are monolingual, bilingual

or multilingual – and this grouping was used during the analysis. Furthermore, as the ‘monolingual’ group remained quite small (N=5), I further divided the group to two main groups: monolinguals and non-monolinguals, which includes the subgroups ‘bilingual’ and ‘multilingual’, and as there were some differences between these two subgroups, these were later compared to one another.

First, I described the groups by count and percentages to portray the groups in terms of their self-categorizations and their language repertoires. After that, I calculated the effect of background factors to the self-categorization by using cross tabulation and Chi square-tests between two groups: monolinguals vs. non-monolinguals and bilinguals vs. multilinguals. Chi square -tests are commonly used for finding the relationships between categorical variables (Larson-Hall 2012). I used the same method of analysis in finding out whether there were differences in participants’ language identities based on the participant’s grade. Furthermore, when all the three groups (monolinguals, bilinguals, multilinguals) were analyzed, Kruskal-Wallis -test was used to find significant differences between the three groups according to their answers to Likert-scale questions. I further compared two subgroups ‘bilinguals’ and ‘multilinguals’ by using Mann-Whitney U-test. As the normal distribution was not assessed, the non-parametric tests such as Mann-Whitney U-test and Kruskal-Wallis-test were used instead of the parametric T-test and ANOVA (Larson-Hall 2012). Kruskal-Wallis is suitable for comparing three groups, as the U-test is used for comparing two groups on continuous variables.

The data received from the open-ended question was analyzed manually, as the answers were relatively short and easy to analyze according to their contents. As such, open-ended questions were analyzed by using content analysis by counting the occurrences of different explanations and by grouping them together to be able to form a categorization of the most common themes. Content analysis is a commonly used method for analyzing the answers on open-ended questions on surveys by coding the data systematically (Friedman 2012, 191). As the majority of respondents gave more than one explanation, and the numbers describe how often a specific explanation was used by the participants, and the examples include only the relevant parts of the answer. Furthermore, the examples are presented in the original form without any corrections. In the next section, I will present the results of this study.

6 Results

In this section, I present the results of my study and answer the research questions set earlier. The aim is to establish a more generalized description of IB-students’ language identity than

in the previous studies in Finland. In this section, I analyze the results in a specific order which follows the research questions set for this study and proceed thematically throughout the section. The sample is divided into groups in two ways based on the research question: primarily, the participants are divided into three groups based on their self-categorization into monolinguals, bilinguals and multilinguals. Furthermore, when relevant, bilinguals and multilinguals are combined to form the group of non-monolinguals to be able to analyze the central differences between monolinguals and those who categorize themselves as either bilinguals or multilinguals. First, I shortly describe the primary characteristics of the three main groups based on the participants' self-categorization into monolinguals, bilinguals and multilinguals. That is, I present the division into these three language identity groups and discuss their language repertoires based on the data gathered.

Secondly, I analyzed the Likert-scale questionnaires quantitatively by using Kruskal-Wallis-tests when analyzing the differences in the agreement with the statements between all three groups (monolinguals, bilinguals and multilinguals), and further using Mann-Whitney U-test to find out whether there are significant differences between bilinguals and multilinguals. This section further discusses the characteristics of the participants' language identities, as the participants' answers on Likert-scale questions and the open-ended question were designed to reveal more about the subjective experiences of bilingualism. That is, in this section, both quantitative and qualitative analysis were conducted.

Next, to answer to the other two research questions set for this study, I present the descriptive statistics and statistical analysis of a few questionnaire items to describe the extent to which participants' differ according to their background and whether there are any differences between three groups of students: pre-IB (group 1), 1st grade (group 2) and 2nd grade (group 3). I compared monolinguals' answers to those of non-monolinguals (bilinguals and multilinguals) and demonstrate the central differences between these two main groups, as well as statistically compared the two subgroups 'bilinguals' and 'multilinguals' with one another.

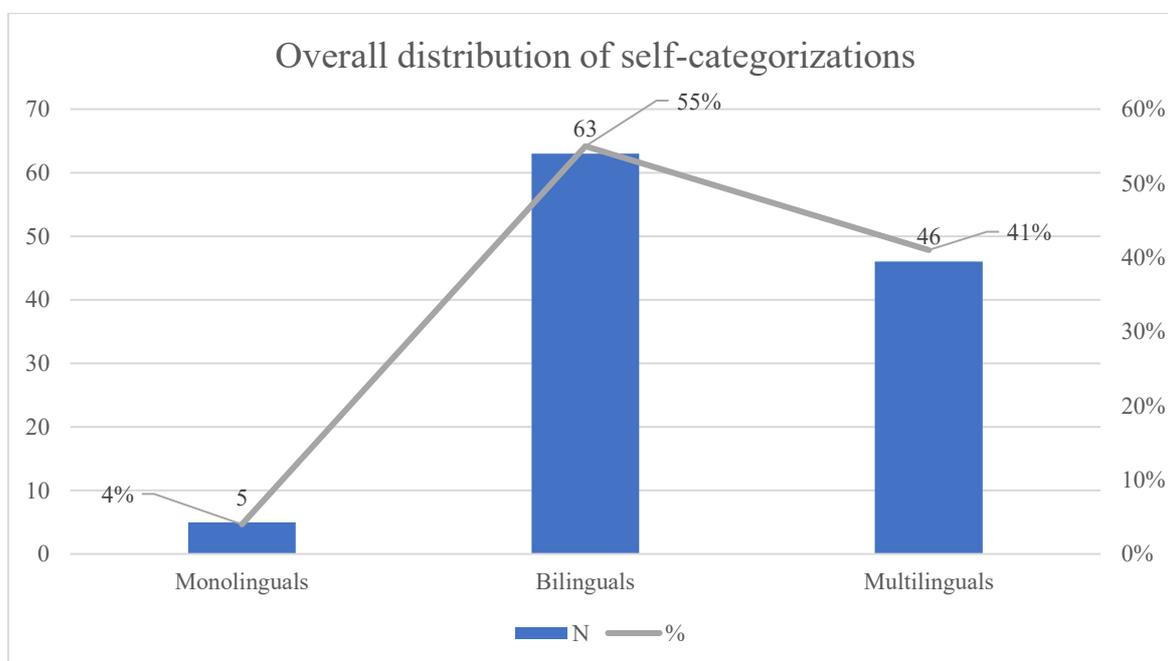
6.1 Language identity of Finnish IB-students: monolinguals and non-monolinguals

In this subsection, the main goal is to discuss the characteristics of three groups of participants based on their self-categorized language identity. First, I present the overall distribution of self-categorizations and then move on to the participants' language repertoires. Furthermore, I present the participants' own explanations of how the number of

languages and the age of acquisition have affected their views of their own language identity based on the answers on the open-ended question.

According to the data received from the questionnaire, most of the participants (N=109/114) describe themselves as non-monolinguals. This means that, when participants were asked to categorize themselves in terms of being mono-, bi- or multilingual, only 5 participants (4%) chose the alternative “monolingual” (see Figure 2). In this study, I chose the term ‘non-monolingual’ to refer to all participants who reported being either bilingual or multilingual. This way, I can investigate the participants who feel ‘non-monolingual’ for any reason as one larger group and further divide it into two subgroups: ‘bilinguals’ and ‘multilinguals’.

Figure 2. Overall distribution of self-categorizations.

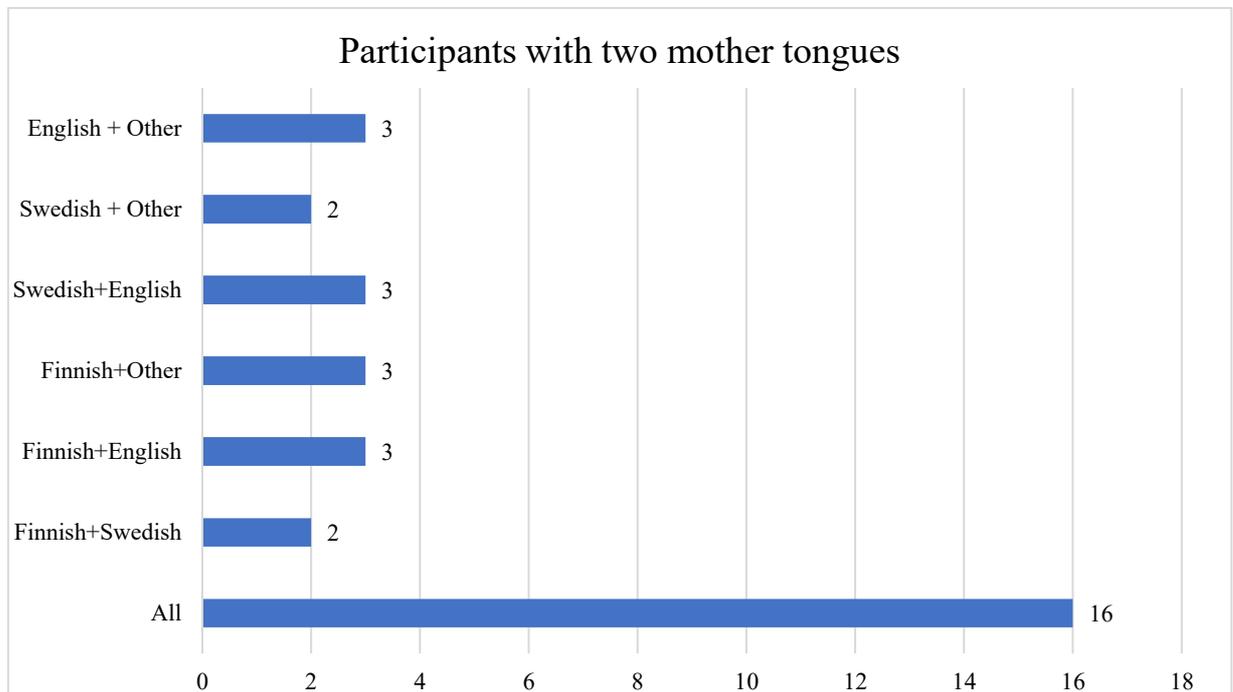


The figure above indicates that among IB-students in Finland, students who feel monolingual are a minority. Furthermore, from the overall number of non-monolinguals as a group including both self-categorized bilinguals and multilinguals (N=109), the option ‘bilingual’ received more answers (N=63; 55%). This means that ‘bilingual’ is the most common way for IB-students to categorize themselves, when given the opportunity to choose between three options (monolingual, bilingual, multilingual). Still, 41% of the participants defined themselves as multilinguals (N=46). The results are different from previous studies conducted in Finland (McCambridge 2007; Kovanen 2011), as in those studies, it was found that IB-students mostly refer to themselves as monolinguals. This could be affected by the

fact that the participants in McCambridge's and Kovanen's studies were all native Finnish-speaking individuals, as in this study, students come from various backgrounds. Furthermore, since the previous studies were conducted a few years ago and English has gained more and more status in Finland since, it could result in multilingual identity formation among younger generations.

From the whole sample (N=114), when asked to tell what their mother tongue is and given the opportunity to choose multiple options, 16 participants chose more than one language as their mother tongue (see Figure 3 below). Still, none chose more than two languages.

Figure 3. Participants with two mother tongues.



The distribution of mother tongues of non-monolinguals is quite equal, and no combination seems more common than others. As can be seen from Figure 3 above, from 16 participants who chose two languages, all combinations received two or three answers. That is, already at the beginning of the questionnaire, 16 participants (14%) clearly leaned towards non-monolingualism by choosing two mother tongues.

By further examining the distribution of the participants' strongest languages, it can be seen that some language combinations are more common than others (see Table 2).

Table 2. The distribution of language competence levels (%).

Level of competence	Overall percentage (%)	Monolinguals (%)	Bilinguals (%)	Multilinguals (%)
My strongest language is...	Finnish (49%)	40	54	43
	English (31%)	20	29	35
	Swedish (12%)	20	14	9
	Other (8%)	20	3	13
My second strongest language is...	English (61%)	60	70	50
	Finnish (22%)	20	24	20
	Swedish (9%)	-	5	15
	Other (8%)	20	2	15
My third strongest language is...	Swedish (42%)	-	49	37
	Other (34%)	20	38	30
	Finnish (18%)	40	13	22
	English (6%)	40	-	11

As shown in Table 2, almost half of the participants reported Finnish as their strongest language, and 31% reported having English as their strongest language. In addition, 61% reported having English as their second strongest language, and Finnish was the second most common choice (22% of the whole group). The most commonly reported third strongest language was Swedish (42%). Surprisingly, none of the monolinguals reported having Swedish as the second or third strongest language.

These results are in line with the expectations as the education is received in Finland, and it is carried out in English. Furthermore, as Swedish is an official language in Finland, it could explain why the third strongest language of Finnish IB-students usually is Swedish. Still, the results would suggest that as the role of English has grown during the last decade and it has received more and more importance in all areas of life (e.g. Taavitsainen and Pahta 2003), the shift towards ‘English as a second language-society’ has progressed to a stage where English is considered as the second language instead of Swedish, and Swedish has been losing its importance at least among IB-students. As such, the language identity of these

participants can be generalized to consist of a combination of two strong languages, Finnish and English.

Of course, among the sample (N=114) there are multiple different combinations, but as can be seen from the table above, Finnish as the strongest and English as the second strongest language is the most common combination. Still, the results show that bilinguals are more homogenous in regards of the choice of languages, and that more variation can be found from the group of multilinguals: percentages show that the differences are not as large as in the group of bilinguals. That is, among bilinguals the second strongest language is English with 70% (N=44), when only 24% (N=15) picked the second most frequent option, Finnish. In opposition, this difference between the two most commonly picked options among multilinguals is smaller: only 50% chose English as their second strongest language, and 20% chose Finnish. This would suggest that in this sample, ‘multilinguals’ is a more heterogenous group regarding their language repertoires. Furthermore, as the most common language identity was non-monolingual, the change towards feelings of bilingualism and multilingualism could be connected to learning other languages and using them in everyday life, which allows these students to choose their own language identity (e.g. Kanno 2003).

Related to the number of mother tongues and language repertoires, when the content analysis of the open-ended questions was conducted, the importance of the age of acquisition and mother tongue(s) was among the most common descriptors of one’s language identity. That is, among monolinguals (N=5) the explanation was used twice, whereas among bilinguals (N=63) it was used by three participants and among multilinguals (N=46) by 4 participants (see Table 3 below).

Table 3. Subjective explanations of bilingualism related to the number of languages and age (N) by the participants.

	Overall (N)	Monolinguals (N)	Bilinguals (N)	Multilinguals (N)
Number of languages	13	0	3	10
Age of acquisition/ mother tongues	9	2	3	4

For these respondents, one of the most important descriptors of bilingualism is the age of acquisition:

- (1) I wasn't born in a multiple languages speaking family (Respondent 39, monolingual).
- (2) I've always seen bilingual as having two mother tongues (Respondent 53, monolingual).

As can be seen from examples 1 and 2, two self-categorized monolinguals explained their feelings of monolingualism as not having been born in a family where more than one language is spoken as a mother tongue. That is, monolinguals tended to explain their monolingualism with the number of mother tongues. In opposition, non-monolinguals tended to explain their language identity by both multiple mother tongues and the age of acquisition:

(3) I have been learning English since I was 4 (Respondent 100, bilingual).

(4) I have spoken both Swedish and Finnish as my first language my whole life (Respondent 58, multilingual).

As example 3 shows, an important descriptor of the participants' bilingualism is the fact that (s)he has started the acquisition of English at the age of 4. That could be seen either as a simultaneous or sequential bilingualism, the limit age for simultaneous bilingualism differs depending on the theorist (e.g. Yip & Matthews 2007; Hakuta 2010). Furthermore, as example 4 indicates, having multiple first languages is seen as an explanation of bilingualism by the participant.

In addition to the frequent combinations of the three strongest languages among the students, similarly one of the most frequent explanations of one's language identity was the number of languages (see Table 3). That is, this explanation was used three times among bilinguals and 10 times among multilinguals (see examples 5 and 6). Monolinguals did not explain their monolingualism with the number of languages.

(5) I speak two languages (Respondent 36, bilingual).

(6) I speak more than four languages (Respondent 12, multilingual).

A number of similar answers were found from the data. Thus, among the respondents, the number of languages in the participants' repertoire was frequently used without further elaboration: the plain explanation of knowing more than one language was enough for them to feel bilingual. Furthermore, the participants who used the number of languages as a descriptor, did not evaluate the level of competence of those languages. The explanations in which levels of competence were evaluated are discussed later in the next section. The results would suggest that among monolinguals, the number of languages is not an important descriptor of their language identities, as bilinguals and multilinguals tend to characterize their language identity more with the plain number of languages they can speak without further elaboration of the level of competence in those languages. Furthermore, the age of acquisition was used as a descriptor equally in all three groups. For these participants, the most important descriptor of bilingualism seems to be having more than one languages as a

mother tongue, which results in either feelings of bilingualism/multilingualism or monolingualism regardless of other factors such as the level of competence or frequency of language use, which are commonly used descriptors of bilingualism (see Table 1 in section 2).

As the most important characteristics of the two main groups have been discussed, I proceed to analyzing further what the language identities are like from the students' subjective point of view. Thus, I present the differences and similarities between language identity groups based on their answers on Likert-scale questions and open-ended question.

6.2 Subjective experiences of bilingualism

To further analyze what the language identities of IB-students are like, I analyzed the differences between language identity groups based on their answers on Likert-scale questions regarding their language proficiency and confidence, identity change, feelings of bilingualism and identification with speaker communities. First, I compared the answers of all three groups to one another with Kruskal-Wallis-test, and then compared the bilinguals' and multilinguals' answers with Mann-Whitney U-test. I gathered the results of the two comparisons and found some differences between the groups. Thus, a few themes were created based on which the results are presented. Accordingly, the qualitative content analysis of the open-ended question describing the participants' own explanations of their language identities is conducted by using the same thematical subsections.

6.2.1 Language competence, language use and others' views of one's bilingualism

In the first question pattern, there were questions related to language competence, language use and how the participants' think others see them in regards of language. After comparing the means of all three groups (monolinguals, bilinguals and multilinguals) and further comparing bilinguals' and multilinguals' answers to one another, I found some significant differences in the answers to questions "In my opinion, my friends consider me as bilingual/multilingual" and "In my opinion, my family considers me as bilingual/multilingual"(see Table 4 below).

Table 4. Language learning and others' views of one's bilingualism/multilingualism.

	Monolinguals		Bilinguals		Multilinguals	
	M	SD	M	SD	M	SD
2. In my opinion, I have reached the highest competence level (reading, writing, speaking and listening) I will be able to achieve in English.	1.80	1.304	2.32	0.877	2.35	0.795
3. In my opinion, I have made progress in learning English during my IB-studies.	3.40	0.894	3.27	0.66	3.37	0.799
4. At school, I usually communicate with others in English also outside classroom.	3.40	0.894	3.13	0.889	3.41	0.748
5. In my opinion, my friends consider me as bilingual/multilingual.	1.40	0.894	3.13	0.852	3.65	0.737
6. In my opinion, my family considers me as bilingual/multilingual.	1.40	0.894	3.10	0.893	3.70	0.662

When asked to evaluate agreement with the statements, among monolinguals the mean was 1.40 in both questions (SD= 0.894), which means disagreement with the statement, whereas bilinguals and multilinguals tended to agree (bilinguals: 5. M= 3.13, SD=0.852 and 6. M=3.10, SD=0.893, multilinguals: 5. M= 3.65, SD=0.737, 6. M= 3.70, SD=0.662). Kruskal-Wallis-test revealed statistical difference between all three groups in both statements ($p < 0.001$ between bilinguals and multilinguals and monolinguals and multilinguals, between monolinguals and bilinguals $p = 0.005$ in question 5 and $p = 0.007$ in question 6). Moreover, according to Mann-Whitney U-test which was used for the comparison of bilinguals' and multilinguals' agreement, statistical difference remained ($p < 0.001$ in both questions, 5. $U = 894$, 6. $U = 828.5$).

These results indicate that monolinguals tend to feel like their friends and family do not consider them as bi- or multilinguals, which could be the result of their own subjective feelings of monolingualism: as they define themselves as monolinguals, they feel that others

do too. Furthermore, multilinguals tended to agree with the statements more strongly than bilinguals. The possible explanation could be similar to that of monolinguals: it could be that multilinguals feel more strongly being non-monolinguals and feel that others view them in the same way. The differences in the means and standard deviation between bilinguals and multilinguals would back that suggestion in a way that the agreement with the statements deviates from the mean more among bilinguals than multilinguals. That could be the result of adopting the identity and the choice of making it visible to others as discussed in section 2 (Wearing 2011, ix): if the individual keeps the bilingual identity to oneself, it could be that the individual feels that others cannot see that identity either.

In addition, there were no statistical difference between groups in the level of agreement with statements 2 and 3: still, it can be seen that monolinguals disagreed with having reached the highest level of competence possible more than non-monolinguals, whereas all three groups agreed with the statement of having made progress in English language learning during IB-studies. The results among monolinguals are in line with the expectation that if one of the most important descriptors of bilingualism among Finnish IB-students is language competence, monolinguals would feel that they have more progress to be made than individuals who already feel bilingual or multilingual. Also, although no significant difference between groups was found regarding statement 4 about the language use outside classroom, all groups tended to agree with the statement (monolinguals: $M= 3.4$, $SD=0.894$, bilinguals: $M= 3.13$, $SD=0.889$, multilinguals: $M= 3.41$, $SD= 0.748$). That is, participants tend to use English also outside classroom, which would make the overall use of English more frequent.

Moreover, similar results were found in the content analysis of the open-ended question. When asked to explain their choice of language identity, two of the most commonly mentioned reasons were the level of competence (very high or high enough for a bilingual), fluency and nativeness (see Table 5 below).

Table 5. Subjective explanations of bilingualism related to language competence and language use (N) by the participants.

	Overall (N)	Monolinguals (N)	Bilinguals (N)	Multilinguals (N)
High or high enough a level of competence	43	1	30	12
Fluency	38	2	17	19

Frequency	29	1	19	9
Nativeness	6	0	4	2

In general, the level of competence was mentioned once among monolinguals, 30 times among bilinguals and 12 times among multilinguals (examples 7 and 8).

(7) I don't know how to speak Finnish so well (Respondent 9, monolingual).

(8) I feel like I would cope to live in another country where they speak English (Respondent 31, bilingual)

Here, in example 7, the monolingual respondent explained his/her monolingualism by a low level of competence in Finnish. In addition, in example 8, the bilingual respondent justifies his/her bilingualism with a level of competence which is at least high enough to be able to live in an English-speaking country. That is, level of competence was commonly used explanation for one's choice of language identity. Still, the accepted level of competence is fully subjective, and it may differ across the sample. Thus, subjectively set limit for a level of competence that is high enough for the individual was often used as an explanation.

Fluency (examples 9, 10 and 11), which can be thought to be a specific skill as a part of language competence, was also used quite often as a separate explanation (monolinguals: N= 2, bilinguals: N= 17, multilinguals: N= 19).

(9) I am fluent in at least two languages (Respondent 43, bilingual).

(10) I don't feel fluent in English (Respondent 39, monolingual)

(11) I can speak English and Finnish fluently and I am OK at speaking, writing and listening to Spanish and I know the basics of Swedish language (Respondent 109, multilingual).

Often when fluency was used as a descriptor of one's language identity, it was used alone as in example 9 above. In example 10, the respondent justifies the choice of being monolingual with not being fluent in English. Furthermore, in some cases, other explanations were used in addition to fluency, as in example 11: the respondent mentions being fluent in two languages, having a moderate level of competence in different language skills in a third, and even knowing the basics of a fourth language. That is, the respondent used two explanations to justify being multilingual.

The third commonly used explanation that is related to language proficiency was nativeness. Among monolinguals 'nativeness' was not mentioned, but bilinguals (N=4) and

multilinguals (N=2) used it as a descriptor of their language identities. Below, examples 12 and 13 include nativeness as a descriptor:

(12) i can speak two language at a native level (Respondent 3, bilingual).

(13) my English skills are that of a native speaker's (Respondent 103, bilingual).

In the data, as in the examples above, nativeness was mentioned as a separate descriptor, without mentioning mother tongues or levels of competence, but it was often used as a synonym for a high level of proficiency. That is, for these participants, one of the most important descriptors of bilingualism is the native-like knowledge of more than one languages. As nativeness was mentioned only 6 times in the data overall, it cannot be viewed as being very descriptive of the participants' subjective experiences of their language identities. The results are the opposite of Pavlenko's (2003) findings: that is, viewing nativeness as the most important descriptor of bilingualism could result in feelings of incapability when the native level is not achieved, and that way result in categorizing oneself as a monolingual. In this sample, nativeness was not used as an important descriptor, and monolinguals did not describe themselves as monolinguals because of non-nativeness. These results suggest that at least in this sample, nativeness is not seen as the ultimate goal or the ultimate descriptor of bilingualism.

As table 5 above indicates, also the frequency of language use was among the most common descriptors of bilingualism. As the participants study in English, it could be claimed that all use English frequently. Furthermore, as mentioned earlier, agreement with the statement 4 suggests that participants use English also outside classroom. Thus, it could be argued that English is a major part of the participants everyday life, and it is not surprising that the frequency of language use is also used as an important descriptor of bilingualism by the participants. Among monolinguals, the frequency was used once, but again, bilinguals (N=19) and multilinguals (N=9) mentioned frequency more often (see examples 14 and 15 below).

(14) I use both English and Finnish in my everyday life (Respondent 19, bilingual).

(15) i use several languages in my day to day life (Respondent 18, multilingual).

The examples indicate the meaning of regular use of more than one language for the participants. That is, overall frequency of language use was mentioned 29 times in the data as being a descriptor of one's language identity. These results are in line with the theories presented earlier in this study, as regular use of more than one language system is often considered as a definition of bilingualism (e.g. Grosjean 2008).

To conclude, statistical difference was found regarding the agreement with statements of how others view the individuals. Monolinguals did not feel like others would see them as bilinguals/multilinguals, and among non-monolinguals, the group of multilinguals agreed with the statement more strongly. Furthermore, although statistical differences were not found, according to the results, all groups felt like they had made progress during their IB-studies, but monolinguals felt that they had more to achieve regarding English language competence. Also, all groups tended to use English even outside classroom. Qualitative analysis revealed similar results: among the most important descriptors of their own language identity were explanations related to language proficiency and the frequency of language use.

6.2.2 Identity change

The second question pattern was related to identity and identity change. The statements were designed to indicate the feelings of language identity change regarding two different languages: Finnish and English. In these statements, one statistical difference between all groups was found.

Kruskal-Wallis-test revealed significant difference between all three groups in the agreement with statement 8 (between monolinguals and bilinguals $p=0.014$, between monolinguals and multilinguals $p= 0.039$). As can be seen from the Table 6 below, monolinguals tended to disagree with the statement “I feel like a different person when I speak English” ($M= 1.60$, $SD= 0.894$).

Table 6. Confidence, identity and identity change during IB-studies.

	Monolinguals		Bilinguals		Multilinguals	
	M	SD	M	SD	M	SD
7. I feel more confident in English than in other languages	2.4	1.517	2.92	0.809	2.98	0.931
8. I feel like a different person when I speak English.	1.6	0.894	2.78	0.941	2.61	0.954
9. I feel that my identity as an English speaker has become stronger during my IB-studies.	2.6	1.517	3.25	0.74	3.22	0.814
12. I feel that my identity as a Finnish speaker has become weaker during my IB-studies.	1.4	0.548	1.92	0,885	2.174	1.141

In opposition non-monolinguals tended to agree with the statement more often, but the distribution among non-monolinguals was not large: bilinguals ($M=2.78$, $SD= 0.941$) agreed with the statement more often than multilinguals ($M= 2.61$, $SD= 0.954$). Still, no significant difference was found when the U-test was conducted between bilinguals and multilinguals. Furthermore, although no significant difference between the groups was found, when asked whether they felt that their identities as English speakers and that of Finnish speakers had become stronger or weaker, monolinguals' answers differ from those of non-monolinguals. Non-monolinguals felt that their English identity had become stronger during IB-studies (bilinguals $M=3.25$, $SD= 0.740$, multilinguals $M=3.22$, $SD= 0.814$) in opposition to monolinguals, who did not agree with the statement as strongly ($M=2.60$), but their answers vary between the group members quite much ($SD= 1.517$). In addition, monolinguals disagreed with the statement about their Finnish identity becoming weaker ($M=1.40$, $SD= 0.548$) more than non-monolinguals. Moreover, in relation to identity, the participants were asked whether they feel more confident when speaking in English than in other languages. No significant differences were found, and the distribution of answers vary in all groups from disagreement to agreement (monolinguals: $M=2.40$, $SD= 1.517$, bilinguals: $M= 2.90$, $SD= 0.809$, multilinguals $M= 2.98$, $SD= 0.931$).

The results suggest that monolinguals in this study possess a moderately stable non-multilingual identity as they do not feel like different persons when they speak in English, but their English identities and Finnish identities have not changed either. In opposition, bilinguals feel different when speaking in English and feel that their English identity has changed, but Finnish identity has not. As multilinguals agreed with the statements less than bilinguals, it would seem that bilinguals possess stronger Finnish identity which has stayed more unaffected during IB-studies than that of multilinguals. It would seem that monolinguals have not created a bilingual identity but possess a strong Finnish identity, as bilinguals and multilinguals have changed their identity towards more bilingual/multilingual identity. Furthermore, as they feel different when speaking in English, it could be a result of a non-monolingual identity: as they have accepted the non-monolingual identity and the two (or more) languages and cultures are isolated entities in individuals, they kind of change roles when speaking in English. That is, when the individuals have adopted the monolingual view of bilingualism (Grosjean 2008), they may feel like they possess two isolated systems which are switched on by using the language. In opposition, if the individual has adopted the bilingual view of bilingualism (ibid.), (s)he may not feel this kind of discrepancy when switching between languages.

Furthermore, the answers on open-ended question indicate similar tendencies in describing bilingualism with identity-related explanations, confidence and code-switching (see Table 7). Here, other affective explanations were grouped together (N=4), and confidence formed its own category as it was mentioned separately in the data (N=3).

Table 7. Subjective explanations of bilingualism related to identity and code-switching (N) by the participants.

	Overall (N)	Monolinguals (N)	Bilinguals (N)	Multilinguals (N)
Affective explanations	4	0	1	3
Code-switching	3	0	2	1
Confidence	3	1	1	1

In relation to the Likert-questionnaire item about confidence, it seems that confidence in speaking multiple languages is used as a descriptor of bilingualism:

(16) I'm not very confident with my English yet (Respondent 111, monolingual).

(17) my mother tongue is Swedish but I can also speak English just as confidently (Respondent 25, multilingual)

That is, being confident is one descriptor of bilingualism, although not very frequently used by the participants. Feelings of confidence may be the result of multiple factors, such as high competence which has been achieved through CLIL-education (e.g. Juan-Garau and Salazar-Noguera 2016). Thus, it could be that participants who have high competence in multiple languages for any reason feel also more confident using them. That is why some participants may have explained their bilingualism as feelings of confidence. Still, this cannot be verified in the light of this data.

Other affective factors, related to any feeling towards language, language community or language identity, were mentioned once by bilinguals and three times by multilinguals (see examples 18 and 19):

(18) I'm also proud of the languages that I am able to speak (Respondent 30, multilingual).

(19) I am a bilingual, because I feel like one (Respondent 32, multilingual).

That is, in the examples above, both participants explained their bilingualism with affective factors: in the first example, the respondent has feelings of pride, whereas in the second example, the respondent just mentions “feeling bilingual”, so (s)he must be one. Affective factors were not used often, but they are as valid explanations as any other: language identity is a subjective feeling of belonging to an ingroup, and emotions towards languages and language communities are acceptable factors in describing bilingualism (see Table 1). For these participants, the subjective experience of bilingualism through emotions is among the most important descriptors of their language identity.

Furthermore, code-switching was used as an explanation of bilingualism (N=3). Again, monolinguals did not mention code-switching, whereas bilinguals (N=2) and multilinguals (N=1) did use it as a descriptor of their language identities.

(20) often when i cant find the word in the other language i always know it on the other (Respondent 5, bilingual)

(21) Due to my bilingual nature, I at times forget the Finnish equivalent of English words, something that is associated with being bilingual (Respondent 90, bilingual).

In these examples, the participants are explaining their use of code-switching as one of their communication strategies. Furthermore, the respondent 90 even mentions that code-switching is common among bilinguals, and that way justifies his/her choice of language identity. As discussed earlier in this study, Bhatia & Ritchie (2012) also mentioned this linguistic alteration for a variety of reasons as a property of bilinguals. Thus, these results prove that also some of the participants of this study feel more bilingual as a result of code-switching, even though it was not very frequently used.

To conclude, monolinguals of this study possess a rather stable language identity that is not affected by either IB-education or English language use. Furthermore, the language identities of non-monolinguals have changed more during IB-studies and they have felt different when speaking in English. This could be the result of individuals’ own views of bilingualism. In addition, qualitative analysis revealed that the non-monolingual participants tend to explain their language identities with different affective factors and the frequent code-switching, whereas confidence was used in all groups as a descriptor. These results strengthen the view that language identity is not a stable construction but changes throughout lifetime, and that emotions and affective factors play a crucial part in one’s language identity construction.

6.2.3 Subjective feelings of bilingualism

The third question pattern included questions of the participants' subjective feelings of bilingualism. The questions were constructed in the same manner, starting with "I feel bilingual when...". With the help of these statements, differences in subjective feelings of bilingualism between groups were found.

First, Kruskal-Wallis revealed significant difference between multilinguals and monolinguals/bilinguals in questions 14 and 15 about interaction with family and friends (see Table 8 below). The differences between monolinguals and multilinguals were in question 14. $p=0.004$ and in question 15. $p=0.01$. This means that, in both questions related to situations of interaction with the nearest others, monolinguals tended to disagree (14. $M=1.8$, $SD=1.304$, 15. $M=2.4$, $SD=0.894$), whereas multilinguals tended to agree (14. $M=3.33$, $SD=0.920$, 15. $M=3.44$, $SD=0.860$). Furthermore, the difference between bilinguals and multilinguals in question 14 was significant ($p<0.001$), whereas in question 15 the difference between monolinguals and bilinguals was significant ($p=0.047$). This would suggest that monolinguals differ more from multilinguals than from bilinguals regarding feelings of bilingualism when interacting with the nearest.

Table 8. Subjective feelings of bilingualism.

	Monolinguals		Bilinguals		Multilinguals	
	M	SD	M	SD	M	SD
14. I feel bilingual when I am interacting with my family.	1.80	1.304	2.54	0.997	3.33	0.920
15. I feel bilingual when I am interacting with my friends.	2.40	0.894	3.25	0.782	3.44	0.860
16. I feel bilingual when I am travelling in English speaking countries.	2.80	1.095	3.29	0.728	3.59	0.858
17. I feel bilingual when I am travelling in non-English speaking countries.	2.80	0.837	2.81	0.931	3.15	0.965
18. I feel bilingual when I am speaking English with a native English speaker.	2.00	1.225	3.08	0.747	3.13	0.980
20. I feel more bilingual now than at the beginning of my studies in IB	2.40	0.894	3.14	0.84	3.04	1.074

Moreover, U-test revealed significant difference between bilinguals and multilinguals in question 14 ($p < 0.001$, $U = 817$). That is, multilinguals agreed more strongly with the statement than bilinguals about feeling bilingual when speaking with family members. This could be the result of situational effects or the nature of language identity: if the interlocutor views the participant as bilingual, the participant may feel more bilingual. Another explanation could be that if the interlocutors are themselves bilinguals, the participant feels more bilingual, when the bilingual language identity is activated. This is related to the bilingual view of bilingualism, where bilingual is viewed as having two intertwined language systems and can use different languages and combinations of them in different situations, with different people, for different purposes (Grosjean 2008; Bhatia and Ritchie 2012). Moreover, it could be the consequence of situational language identity: the language identity is created for specific situations, and it is not present in other situations (Korostelina 2007). That could result in feelings of bilingualism in one situation, but not in another.

Secondly, agreement with the statements related to travelling also revealed significant differences between the groups. In question 16 “I feel bilingual when I am travelling in English speaking countries”, multilinguals tended to agree strongly ($M = 3.59$, $SD = 0.858$), as monolinguals agreement varied more between disagreement and agreement ($M = 2.8$, $SD = 1.095$). Kruskal-Wallis revealed significant difference in the agreement between monolinguals and multilinguals ($p = 0.04$) and bilinguals and multilinguals ($p = 0.004$). Similar results were received when U-test was used between bilinguals and multilinguals ($p = 0.003$, $U = 1023$). Furthermore, significant difference between bilinguals and multilinguals was found in the agreement with the statement 17 about feeling bilingual when travelling in non-English speaking countries ($p = 0.039$, $U = 1129.5$). As mentioned above, situational identities could be the reason for these results, as the bilingual identities could be attached to concrete situations, such as travelling in different countries where the participants meet either English-speaking or non-English speaking people. As such, individuals may have created an identity where they identify themselves into belonging to non-English speaking ingroup or English-speaking ingroup, which further describes how they compare themselves with the outgroup in specific situations (Tajfel and Turner 1979). Still, bilinguals seemed to feel less bilingual when travelling in non-English speaking countries ($M = 2.81$, $SD = 0.931$) than multilinguals ($M = 3.15$, $SD = 0.965$), although both groups felt less bilingual in non-English speaking countries in comparison to travelling in English-speaking countries. The reason for these results could be an interesting research question for further studies on the subject.

Even though no significance was found in other questions of this question pattern, interesting results were found in questions 18 and 20 that should be discussed here. In question 18, the participants were asked to evaluate how bilingual they feel when they are speaking in English with a native English speaker. Monolinguals tended to disagree ($M=2.00$, $SD=1.225$), whereas bilinguals tended to agree ($M=3.08$, $SD=0.747$). Furthermore, multilinguals tended to agree even more strongly ($M=3.13$, $SD=0.98$). This would suggest that when the participants speak with an individual who they know is a native English speaker, monolinguals feel monolingual, and non-monolinguals tend to feel bilingual. That is, they may feel competent enough to describe themselves as belonging in a similar ingroup than the native speaker, as monolinguals may not feel competent enough in comparison with the native speaker (e.g. Hakuta 2010).

Moreover, when the participants were asked whether they feel more bilingual now than at the beginning of their studies on IB, monolinguals tended to disagree more often ($M=2.4$, $SD=0.894$) with the statement than non-monolinguals (bilinguals $M=3.14$, $SD=0.84$ and multilinguals $M=3.04$, $SD=1.074$). Furthermore, as bilinguals agreed with the statement more and their answers had less variation, it could be suggested that their bilingual identity has developed more than that of multilinguals. It could be a result of a variety reasons, but as it could be seen from the results presented throughout the results-section, it would seem that multilinguals have already formed a more stable language identity and that is why they have stronger agreement with the statements considering subjective feelings of bilingualism and weaker agreement with the statements considering identity change than bilinguals. Thus, it could be hypothesized that the participants who at the moment self-categorize themselves as bilinguals, have lately formed their bilingual identity or they are still constructing it, and that is why they do not feel like non-monolinguals as strongly as self-categorized multilinguals. That is, they can still be processing and forming their language identities, so they have not achieved the stable phase yet (Korostelina 2007). Of course, this cannot be verified by the length of this study.

To conclude, non-monolinguals have stronger subjective feelings of bilingualism than monolinguals. Still, it would seem that multilinguals tend to feel bilingualism more strongly than bilinguals. This could be because of the ongoing language identity formation process of bilinguals, since bilinguals felt more progress in language identity formation during IB-studies than multilinguals. Moreover, the differences in the agreement with the statements between bilinguals and multilinguals could be the result of the form of language identity:

whether it is still being constructed, or whether it is situational or stable after it has been formed.

6.2.4 Identification with language communities

In the last question pattern, there were six statements which were designed to measure the degree of identification with language communities. As in the previous patterns, also in these questions some significant differences were found when the means of all groups were compared to one another. The most important results are presented in Table 9 below.

Table 9. Identification with different language communities.

	Monolinguals		Bilinguals		Multilinguals	
	M	SD	M	SD	M	SD
22. I identify myself with native speakers of English.	2.00	1.414	2.11	1.002	2.33	0,871
23. I identify myself with non-native speakers of English.	2.40	0.894	2.97	0.761	2.89	0.823
24. I identify myself with monolingual Finnish speakers	1.60	0.894	1.91	0.928	1.717	0.935
25. I identify myself with bilingual English speakers.	2.00	1.000	3.16	0.677	2.739	0.88
26. I identify myself with multilingual speakers.	1.80	0.837	2.46	0.947	3.522	0.547

On one hand, as can be seen from the table above, none of the groups identified themselves with native speakers of English (monolinguals: M= 2.00, SD= 1.414, bilinguals: M= 2.11, SD=1.002, multilinguals: M= 2.33, SD= 0.871). Still, as the deviation in the groups of monolinguals and bilinguals are higher, it would seem that there is quite a lot of variation in the answers of those groups. As such, the results would suggest that the participants view native speakers of English different from themselves, which could be because of a variety

of reasons: among them could be viewing native-like competence as a synonym for extremely high language proficiency and as the descriptor of a native speaker of English, and therefore not identifying oneself with the group (Hakuta 2010; Kanno 2003). Furthermore, they do not seem to identify themselves (at least very strongly) either with non-native speakers of English or monolingual speakers of Finnish (see Table 9). These results suggest that even the self-categorized monolinguals do not identify themselves as being a monolingual Finnish speakers, and that all the participants may feel feelings of belonging to a language community that is somewhat different from both native and non-native speakers of English.

On the other hand, significant differences were found in the agreement with statements 25 and 26 that ask whether the participants identify themselves with bilingual English speakers or with multilingual speakers. What was found out was that bilinguals tended to identify themselves more strongly with bilingual speakers of English ($M=3.16$, $SD=0.677$) than multilinguals ($M=2.74$, $SD=0.880$). In opposition, multilinguals tended to identify themselves with multilingual speakers clearly more strongly ($M=3.52$, $SD=0.547$) than bilinguals ($M=2.46$, $SD=0.947$). Furthermore, monolinguals did not identify themselves with either of the speaker communities. Kruskal-Wallis revealed significant differences in the identification with bilingual speakers of English between monolinguals/bilinguals ($p=0.009$) and bilinguals/multilinguals were found ($p=0.010$). Again, U-test revealed similar results between bilinguals and multilinguals ($p=0.010$, $U=1062$). Moreover, as U-test revealed difference between bilinguals and multilinguals in identification with multilinguals ($p<0.001$, $U=556$), also Kruskal-Wallis-test backed the results: the difference of both monolinguals vs. multilinguals and bilinguals vs. multilinguals was significant ($p<0.001$).

The results would suggest that the monolingual participants do not identify themselves very strongly with the existing language communities. Accordingly, the results indicate the tendency of bilinguals to identify themselves with bilingual speakers of English, whereas multilinguals tend to identify themselves with multilinguals. It would seem that the questionnaire did not offer an appropriate language community for monolinguals to identify themselves to. These results suggest that monolinguals may have a hybrid identity (Kanno 2003). The results are somewhat in line with different theorists' propositions (e.g. Tajfel and Turner 1979; Grosjean 2008) of identity negotiation and language identity as being reflecting of one's subjective feeling of belonging to a language community. Now that the results of subjective experiences of bilingualism have been discussed, I move on to discussing the effect of various background factors on language identity.

6.3 The effect of background factors

As one of my research questions was the effect of background factors on language identity, I conducted cross tabulation and chi square -tests to see whether background factors were actually related to participants' self-categorizations. Again, I used groups 'monolingual' and 'non-monolingual' which included bilinguals and multilinguals and then compared bilinguals and multilinguals as separate groups in order to see whether any of the background factors were related to one's self-categorization. Background factors were gathered at the beginning of the questionnaire and here I chose to analyze age of acquisition, the amount of time spent abroad in English-speaking countries or non-English speaking countries, and the amount of English-speaking education (by coding the answers according to first level of English-speaking education received) , and the number of languages used at home with family members.

First, after running the cross tabulation and Fisher's exact test of the first two categories (monolinguals and non-monolinguals), I found out that in my sample, there are no significant effects of background factors on the participants' decision between options 'monolingual' or 'bilingual/multilingual' (see Table 10 below, column titled 'monolinguals/non-monolinguals'). This could be the result of rather unequal distribution of participants into the two categories, as monolinguals are a minority in this study. As such, no significant dependence between the variables could be established.

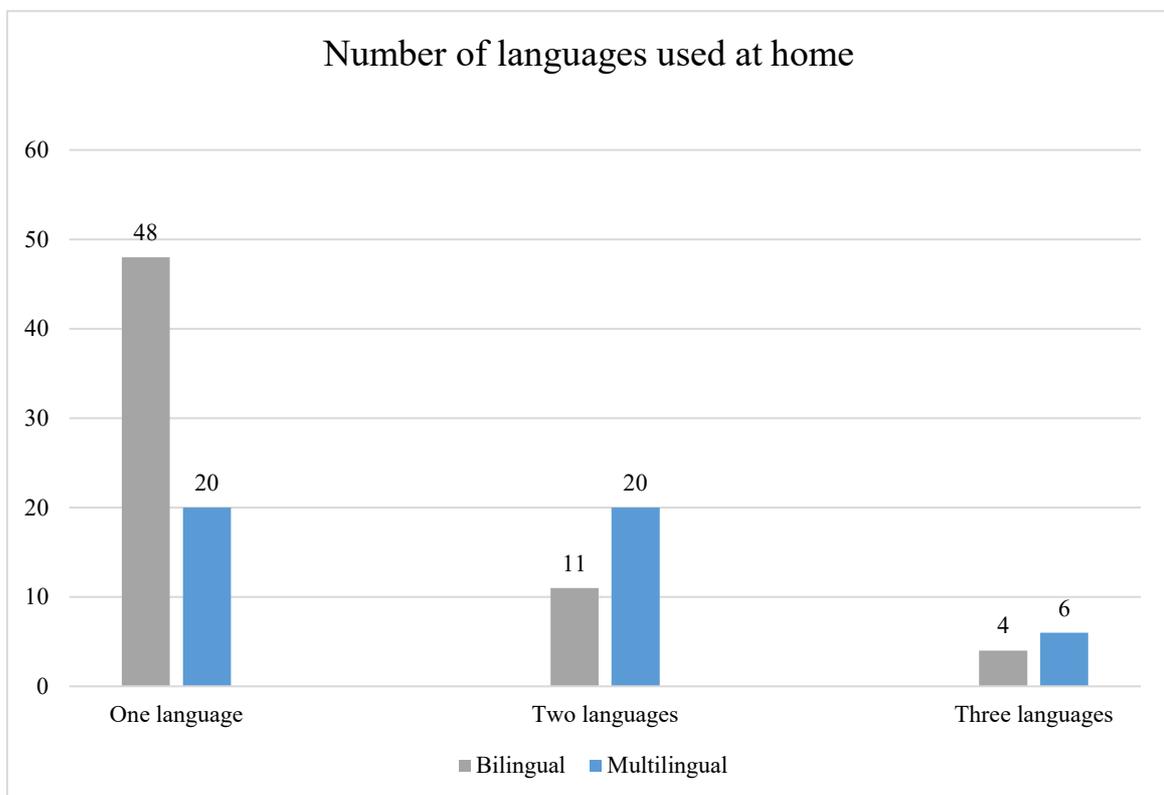
Table 10. The relationship between self-categorization and background (p).

	Monolinguals/non-monolinguals (p)	Bilinguals/multilinguals (p)
Number of languages used at home	1.000	0.002
The first level of English-speaking education	0.152	0.042*
Age of acquisition	0.597	0.118
Living in English-speaking country	0.701	0.431
Living abroad (non-English-speaking)	0.593	0.621

* 4 cells (40%) have the expected count less than 5.

Secondly, I compared self-categorized bilinguals and multilinguals together in relation to background variables (Table 10 above, column titled ‘bilinguals/multilinguals’) and found out that one of the background factors had effects to participants’ self-categorization. Although the significance of the relationship between education and self-categorization is just below the limit of significance ($p < 0.05$), it is not a significant result as 4 cells had the expected count less than 5, which means that the test did not meet the reliability criteria. That is, one of the five background factors were related to participants’ choice between bilingual or multilingual identities: the reported number of languages used at home with other family members. Below, Figure 4 shows the distribution of the number of languages by self-reported bilinguals and multilinguals.

Figure 4. The number of languages used at home with other family members (N).



As can be seen from the figure above, most of the bilinguals use only one language at home (N=48, 76% of bilinguals). Furthermore, 17 % of bilinguals (N=11) reported using two languages, whereas multilinguals’ reports of the number of languages was equal between one and two languages (both N=20, 43% of multilinguals). No significant differences in reports of using three languages by bilinguals and multilinguals when looking at the count and percentages. In addition, Chi square -tests revealed significant dependence between self-categorization and the number of languages used ($\chi^2 = 12.187$; $df = 2$; $p = 0.002$). The results

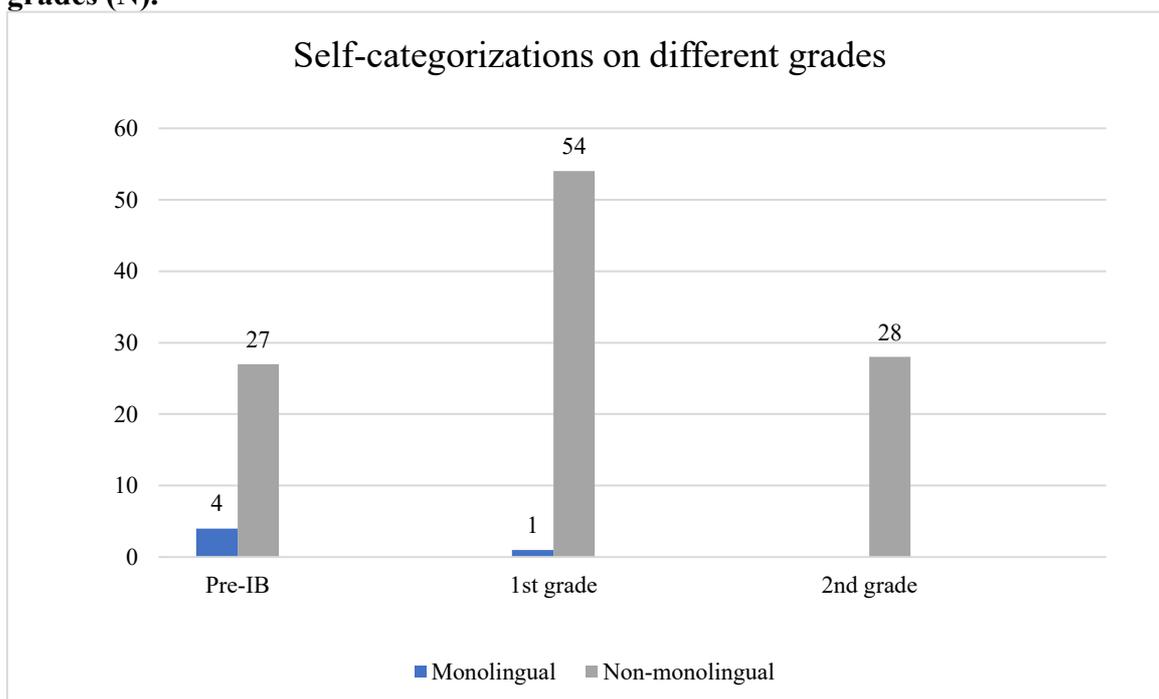
would suggest that the number of languages used at home with other family members would have an effect on the decision between bilingual and multilingual identity.

To conclude, when comparing monolinguals and non-monolinguals (including both bilinguals and multilinguals) no significant result were found when the effects of background factors were analyzed. This could be the result of the unequal group sizes, as monolinguals were a minority in this data. Furthermore, when comparing the effect of background factors on bilinguals and multilinguals, one significant result was found: the number of languages used at home with other family members is related to the choice between categorizing oneself as bilingual or multilingual. The results are similar to theories of bilingualism, as using multiple languages at home suggests that bilingualism is active rather than passive (Hakuta 2010). Moreover, multiple regularly used languages also fits into the definition of bilingualism (Grosjean 2008).

6.4 The differences in self-categorizations between grades

To be able to answer the third research question, I conducted two different Chi square - tests. First, I analyzed the relationship of the participant's grade and the choice between monolingual or non-monolingual identity (by using Fisher's exact test), and after that I continued with analyzing the relationship between grade and the choice between bilingual and multilingual identity. After analyzing the test results, I found out that grade did not have significant effects ($p = 0.054$) on the choice between monolingual and non-monolingual identities (see Figure 5).

Figure 5. Self-categorization of the monolinguals and non-monolinguals on different grades (N).



As can be seen from the figure above, from the participants who chose monolingual identity (N=5), four (80% of the monolinguals) were on pre-IB level at the moment of answering the questionnaire, and one was on the 1st grade of Diploma Programme. Furthermore, no 2nd graders self-categorized themselves as monolinguals. It could be suggested (even though no statistical significance was found in this study) that when the program proceeds and the students receive more English-speaking education, students who at the beginning may have identified themselves as monolinguals, change their language identity. This suggestion is in line with the theories of identity: as discussed in section 2, identity is not stable and unchangeable construction, but it is a dynamic system, and identity formation is an ongoing process throughout lifetime. Still, this change cannot be verified by using this data, as my study is not longitudinal, and my sample is rather uneven in relation to the juxtaposition of monolinguals and non-monolinguals participants.

Secondly, no significant relationship was found between grade and self-categorization into bilinguals and multilinguals either. After running Chi square test between the variables, the results ($\chi^2=3.676$, $df = 2$, $p = 0.159$) show that the relationship between grade and self-categorization is not significant. The distribution of self-categorizations of bilinguals and multilinguals on different grades can be seen from the Table 11 below.

Table 11. Self-categorizations of bilinguals and multilinguals on different grades.

			Bilingual	Multilingual	Total
Grade	Pre-IB	N	18	9	27
		%	66.7%	33.3%	100.0%
	1st grade	N	33	21	54
		%	61.1%	38.9%	100.0%
	2nd grade	N	12	16	28
		%	42.9%	57.1%	100.0%
Total		N	63	46	109
		%	57.8%	42.2%	100.0%

As the table above shows, bilinguals make the majority of pre-IB students and 1st graders, but the situation is the opposite among 2nd graders. Still, the differences in count and percentages between self-categorized bilinguals and monolinguals are the largest on pre-IB level, and the smallest on 2nd grade. Still, no indications of the effect of grade cannot be affirmed, as no statistical significance was found and, as shown in table 3, the group of 1st graders is more represented than the other two groups of students.

7 Conclusion

The main goal of this study was to create a more generalized view of what the Finnish IB-students' language identities are like, and whether background or grade has any effect on the participants' self-categorization. In this study, the participants were divided in two ways according to their self-categorization: first, the participants were divided into monolinguals and non-monolinguals (consisting of bilinguals and multilinguals), and then into three separate groups (monolinguals, bilinguals, multilinguals). The differences were found by examining the differences in the agreement with statements related to subjective feelings of bilingualism, identity and identity change, perception of one's own language competence and others' views of oneself, and identification with language communities. Furthermore, the participants' subjective experiences of their language identities were analyzed.

When looking at the characteristics of the sample (N=114), it seemed that monolinguals were a minority (N=5), as bilinguals (N=63) were a slightly bigger group by count than multilinguals (N=46). Moreover, 16 participants reported having two mother tongues, and that way leaning towards non-monolingualism. Still, none reported having more than two mother tongues. When examining the participants' language repertoires, the results revealed that the most common combination of the two strongest languages was Finnish-English. Swedish tended to be the third strongest language. Furthermore, the age of acquisition, multiple mother tongues and the overall number of languages spoken by the individual were seen as important descriptors of bilingualism according to the participants' answers on the open-ended question. In light of these results, it could be claimed that Finnish IB-students tend to categorize themselves as non-monolinguals, and the number of mother tongues and language repertoires can be seen as predictive factors in the choice between different language identities. These results are different from previous studies conducted in Finland, as in McCambridge's (2007) and Kovanen's (2011) studies, the interviewees categorized themselves as monolinguals. As today the status of English is even higher in Finland, and it has been suggested that Swedish will lose its importance and English will become the second language of the society instead of being a foreign language (e.g. Taavitsainen and Pahta 2003). The results are in line with the expectation of language repertoire changes in the society, as the participants' second strongest language was English.

Subjective experiences of bilingualism were analyzed by examining the participants' answers on Likert-scale questions that formed 4 question patterns with different themes. In the quantitative analysis between the groups, statistical differences were found. The results show that monolinguals tend to feel like their family and friends do not see them as bilinguals, whereas non-monolinguals felt like others viewed them as bilinguals. This could be the result of their own language identity being shared with others, and that way their subjective experience may affect the way they feel others see them, too (Korostelina 2007). Furthermore, as monolinguals tended to feel like they still had more to achieve in English language learning than non-monolinguals, and the level of competence was a common descriptor of bilingualism in all groups, language proficiency may have a crucial role in students' self-categorization.

In addition to level of competence, also fluency, nativeness and frequency of language use were used as descriptors of bilingualism by the participants. These descriptors are in line with theories of bilingualism (e.g. Grosjean 2008; Hakuta 2010; Bhatia and Ritchie 2012). Thus, in the future, encouraging students in their language learning, language use and

language identity formation process, students could gain more confidence and that way Finnish students may feel more confident in categorizing themselves as bilinguals. As regular use of language is mentioned in various definitions of bilingualism (such as Grosjean 2008 and Hakuta 2010), and is also a frequently mentioned descriptor of bilingualism, the analysis of background factors revealed that the number of languages used at home with other family members has a significant effect on language identity. That is, the more languages participants used at home, the more likely they categorized themselves as non-monolinguals. After examining the importance of other background factors and the effect of grade on self-categorization, only the number of languages used at home remained significant.

Moreover, the results indicate that monolinguals of this study had a relatively stable identity that has not changed during IB-studies, whereas bilinguals and multilinguals felt the opposite. Furthermore, non-monolinguals reported feeling different when speaking in English than in other languages. That would suggest the presence of monolingual view of bilingualism (Grosjean 2008) among non-monolinguals: that is, feeling like a different person may be the result of identity discrepancy in switching between languages. Still, switching between languages (that is, code-switching) was seen as a property of bilinguals and one important descriptor of bilingualism. In addition to descriptors mentioned, affective explanations were found important, too. That is, participants used affective explanations, such as emotions and confidence. Still, no difference between groups was found in confidence in using English. In light of the results, it is important to accept code-switching as a communication strategy used in and outside classroom, as it is a part of their identity construction, and a strategy to present themselves to others. Also, as mentioned earlier, confidence and positive emotions are a crucial part of one's language learning and language identity formation process, which should be taken into consideration in IB-education, too: that is, IB-students seem to need encouraging in order to gain more confidence.

When looking at identification with language communities, it was found that monolinguals did not identify themselves with any of the offered language communities, which would suggest that monolinguals have created a hybrid identity, which is a combination of different language identities. Accordingly, bilinguals tended to identify themselves with bilingual speakers of English and multilinguals identified themselves the most strongly with multilingual speakers. These results indicate that one of the most important factors in language identity formation is the process of identification with an ingroup (Tajfel and Turner 1979).

The limitations of this study are related to the representativeness, as these results are not fully generalizable as sample size is rather small and uneven. That is, 1st graders and non-monolinguals were more represented in the sample, and the group of monolinguals remained small. Therefore, for example, the effects of background factors and grade on language identity could be further analyzed by gathering an even larger data with more equal distribution of students from all three grades and try to find out whether there actually is significance between grade and the students' self-categorization. Furthermore, monolinguals could be an interesting object of study in the future, since in this study, the group of monolinguals remained small and no generalizations can be made. Also, the results raise a question about which language community do the monolinguals of this study identify themselves to and why they did not feel like belonging to any of the communities provided by the questionnaire. Moreover, as identity is a changing construct by nature, identity change of IB-students should be studied more by conducting a longitudinal study.

Overall, the results indicate that subjective experiences of bilingualism and multilingualism are stronger in the group of non-monolinguals, and they used a variety of reasons to justify their bilingualism. In opposition, the results suggest that monolinguals may have created a hybrid identity and explain their monolingualism with the lack of proficiency. One important factor in explaining the choice of language identity group seemed to be the regularity of language use, since it was among the most common descriptors used by the participants, and a significance was found in quantitative analysis. Since language identity is how individuals see themselves as language users in relation to others, it is an important part of individuals' identity, and as identity is crucial part of human's psychology, it should be taken into consideration in education too. That is, in the future, in all levels of education, students should be even more encouraged in language learning and use, as it seems that confidence in language use is a crucial part of one's language learning and language identity formation. Although confidence was not the most common explanation used by the participants, it seems that the most important results of this study are related to confidence and competence in a fundamental way. Thus, the formation of students' language identity is a complex process with multiple factors affecting it. As students of any group did not feel very confident in English, their confidence should be promoted more beside focusing on improving language competence.

References

- Bhatia, Tej K., and William C. Ritchie. 2012. *The Handbook of Bilingualism and Multilingualism*. Second edition. Chichester, West Sussex, UK: Wiley-Blackwell. www.b-ok.org
- Breeze, Ruth, Carmen Llamas Saíz, Concepción Martínez Pasamar, and Cristina Taberero Sala, eds. *Integration of Theory and Practice in CLIL. Integration of Theory and Practice in CLIL*. New York: BRILL, 2014. ProQuest Ebook Central. Accessed 8 July 2020.
- Bucholtz, Mary and Kira Hall. 2009. "Locating Identity in Language" In *Language and Identities*, edited by Carmen Llamas and Dominic Watt, 18–28. Edinburgh: Edinburgh University Press. ProQuest Ebook Central. Accessed 19 May 2020.
- Chen, Sylvia Xiaohua. 2015. "Toward a Social Psychology of Bilingualism and Biculturalism." *Asian Journal of Social Psychology* 18, 1: 1–11. EBSCOhost. Accessed 4 September 2020.
- Coyle, Do. 2008. "Content and Language Integrated Learning: Towards a Connected Research Agenda for CLIL Pedagogies." *International journal of bilingual education and bilingualism* 10, no. 5: 543–562. DOI: 10.2167/beb459.0. Accessed 6 August 2020.
- De Jong, Nivja H. 2018. "Fluency in Second Language Testing: Insights From Different Disciplines." *Language Assessment Quarterly* 15, no. 3: 237– 54, DOI: 10.1080/15434303.2018.1477780.
- Dörnyei, Zoltán and Kata Csizér. 2012. "How to Design and Analyze Surveys in Second Language Acquisition Research" In *Research Methods in Second Language Acquisition: A Practical Guide.*, 1. Aufl., edited by Mackey, Alison and Susan M. Gass, 74-93. Wiley-Blackwell. EBSCOhost. Accessed 15 September 2020.
- Dörnyei, Zoltán, and Tatsuya Taguchi. 2009. *Questionnaires in Second Language Research: Construction, Administration, and Processing*. New York: Routledge. EBSCOhost. Accessed 8 September 2020.
- Edwards, John. 2009. *Language and Identity: An introduction*. Cambridge: Cambridge University Press. ProQuest Ebook Central. Accessed 18 April 2020.
- Erikson, Erik H. 1956. "The Problem of Ego Identity." *Journal of the American Psychoanalytic Association* 4, no. 1: 56–121. DOI: 10.1177/000306515600400104. Accessed 18 April 2020.
- Evans, David. 2018. *Language, Identity and Symbolic Culture*. London, UK: Bloomsbury Academic. <http://search.ebscohost.com.ezproxy.utu.fi/login.aspx?direct=true&db=nlebk&AN=1731674&site=ehost-live>. Accessed 18 April 2020.
- Friedman, Debra A. 2012. "How to Collect and Analyze Qualitative Data". In *Research Methods in Second Language Acquisition: A Practical Guide.*, 1. Aufl., edited by Mackey, Alison and Susan M. Gass, 180-200. Wiley-Blackwell. EBSCOhost. Accessed 18 October 2020.
- Grosjean, François. 2008. *Studying Bilinguals*. Oxford University Press. ProQuest Ebook Central. Accessed 28 May 2020.
- Grosjean, François. 2010. *Bilingual: Life and Reality*. Cambridge: Harvard University Press. De Gruyter Mouton. Accessed 3 September 2020.
- Grosjean, François, Ping Li, and Ellen Bialystok. 2013. *The Psycholinguistics of Bilingualism*. Hoboken: Wiley-Blackwell. EBSCOhost. Accessed 8 September 2020.
- Hakuta, K. 2010. "Bilingualism." In *Encyclopedia of Neuroscience*, 173–178. Elsevier Ltd., <https://doi.org/10.1016/B978-008045046-9.01877-5>. Accessed 28 May 2020.
- Halbach, Ana. 2014. "Teaching (in) the foreign language in a CLIL context: Towards a new approach". In *Integration of Theory and Practice in CLIL. Integration of Theory and Practice in CLIL*, edited by Breeze, Ruth, Carmen Llamas Saíz, Concepción Martínez

- Pasamar, and Cristina Taberero Sala, 1-14. New York: BRILL. ProQuest Ebook Central. Accessed 8 July 2020.
- Hamers, Josiane F., and Michel H. A. Blanc. 2000. *Bilinguality and Bilingualism*. Cambridge: Cambridge University Press. ProQuest Ebook Central. Accessed 8 September 2020.
- Higgins, Christina. 2011. "The Formation of L2 Selves in a Globalizing World" In *Identity Formation in Globalizing Contexts. Language Learning in the New Millennium*, edited by Christina Higgins, 1–18. Berlin, Boston: De Gruyter Mouton. Accessed 12 August 2020. <https://www-degruyter-com.ezproxy.utu.fi/view/product/179368>.
- Helsingin Suomalainen Yhteiskoulu. URL: <https://syk.fi/about-ib-at-syk/>. Accessed 14 August 2020.
- International Baccalaureate Organization. URL: www.ibo.org. Accessed 10 August 2020.
- Jenkins, Jennifer. 2007. *English As a Lingua Franca: Attitude and Identity*. Oxford: Oxford University Press. Accessed 19 May 2020.
- Juan-Garau, Maria, and Joana Salazar-Noguera. 2016. *Content-Based Language Learning in Multilingual Educational Environments*. Cham: Springer. EBSCOhost. Accessed 6 August 2020.
- Kanno, Yasuko. 2003. *Negotiating Bilingual and Bicultural Identities : Japanese Returnees Betwixt Two Worlds*. Mahwah: Routledge. ProQuest Ebook Central. Accessed 19 May 2020.
- Kasurinen, Leena. 2019. "Definitions and Subjective Experience of Bilingualism of Finnish University Students of English". MA Thesis.
- Korostelina, Karina V. 2007. *Social Identity and Conflict ; Structures, Dynamics and Implications* . Basingstoke: Palgrave Macmillan. Accessed 18 April 2020
- Kovanen, Heidi. 2011. "How CLIL-Classroom Students See Themselves as Learners of English". MA Thesis. <https://jyx.jyu.fi/bitstream/handle/123456789/27281/1/URN%3ANBN%3Afi%3Aju-2011071111136.pdf>. Accessed 10 September 2020.
- Larson-Hall, Jenifer. 2012. "How to Run Statistical Analyses". In *Research Methods in Second Language Acquisition: A Practical Guide.*, 1. Aufl., edited by Mackey, Alison and Susan M. Gass, 245-274. Wiley-Blackwell. EBSCOhost. Accessed in 18 October 2020.
- Leppänen, Sirpa, and Tarja Nikula. 2007. "Diverse Uses of English in Finnish Society: Discourse-Pragmatic Insights into Media, Educational and Business Contexts." *Multilingua* 26, no. 4: 333–380. EBSCOhost. Accessed 11 September 2020.
- McCambridge, Laura. 2007. "Linguistic and Cultural Identity: Finns Who Have Been Through English Immersion Education". MA Thesis. <https://jyx.jyu.fi/handle/123456789/40552>. Accessed 10 September 2020.
- Nikula, Tarja. 2007. "Speaking English in Finnish content-based classrooms". *World Englishes*, 26 (2), 206-223. DOI: 10.1111/j.1467-971X.2007.00502.x. Accessed 10 September 2020.
- Norton, Bonny. 2000. *Identity and Language Learning: Gender, Ethnicity and Educational Change*. Harlow: Longman.
- Norton Peirce, Bonny. 1995. "Social Identity, Investment, and Language Learning." *Teachers of English to Speakers of Other Languages Quarterly* 29, no. 1: 9–31. Accessed 12 August 2020. www.jstor.org/stable/3587803.
- Luna, David, Torsten Ringberg, Laura A. Peracchio. 2008. "One Individual, Two Identities: Frame Switching Among Biculturals." *Journal of Consumer Research* 35, no. 2: 279–293. doi:10.1086/586914. Accessed 15 June 2020.
- Pappa, Sotiria. 2018. "You've Got the Color, but You Don't Have the Shades": Primary Education CLIL Teachers' Identity Negotiation Within the Finnish Context. Jyväskylä:

- University of Jyväskylä, <https://jyx.jyu.fi/handle/123456789/58586>. Accessed 9 September 2020.
- Pavlenko, Aneta. 2003. "I Never Knew I Was a Bilingual": Reimagining Teacher Identities in TESOL. *Journal of Language, Identity & Education* 2, no. 4: 251–68. Accessed 12 August 2020. EBSCOhost.
- Pérez Cañado and María Luisa. 2018. "The Effects of CLIL on L1 and Content Learning: Updated Empirical Evidence from Monolingual Contexts." *Learning and instruction* 57: 18–33. DOI: 10.1016/j.learninstruc.2017.12.002. Accessed 6 August 2020.
- Pérez Cañado, María Luisa, and Nina Karen Lancaster. 2017. "The Effects of CLIL on Oral Comprehension and Production: a Longitudinal Case Study." *Language, culture, and curriculum* 30, no. 3: 300–316. EBSCOhost. Accessed 6 August 2020.
- Pyykkö, Riitta. 2017. "Monikielisyys vahvuudeksi – Selvitys kansallisen kielivarannon tilasta ja kehittämistarpeista" [Multilingualism into a strength. A report of the status and levels of language competences in Finland.]. Helsinki: Publications of the Ministry of Education and Culture, Finland 51.
- Ringberg, Torsten V, David Luna, Markus Reihlen, and Laura A Peracchio. 2010. "Bicultural-Bilinguals: The Effect of Cultural Frame Switching on Translation Equivalence." *International Journal of Cross Cultural Management* 10, no. 1: 77–92. <https://doi.org/10.1177/1470595809359585>. Accessed 15 June 2020.
- Roiha, Anssi. 2019. "Investigating Former Pupils' Experiences and Perceptions of CLIL in Finland: a Retrospective Analysis." *Nordic journal of studies in educational policy* 5, no. 2: 92–103. <https://jyx.jyu.fi/handle/123456789/66713#>. Accessed 8 September 2020.
- Saastamoinen, Mikko. 2006. "Minuus ja identiteetti tutkimuksen haasteina" [Challenges in Studying The Self and Identity]. In *Minuus ja Identiteetti* [The Self and Identity], edited by Pertti Rautiainen, and Mikko Saastamoinen, 171–80. Tampere: Tampereen Yliopistopaino Oy. Accessed 19 May 2020.
- Surmont, Jill, Esli Struys, Maurits Van Den Noort, and Piet Van de Craen. 2016. "The Effects of CLIL on Mathematical Content Learning: A Longitudinal Study." *Studies in Second Language Learning and Teaching* VI, no. 2: 319–337. URL: http://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight-bdc8e90d-5a26-487e-9a12-47ba41a539c2/c/SSLLT_6_2_319-337_Surmont_Struys_Van_Den_Noort_Van_De_Craen.pdf. Accessed 6 August 2020.
- Taavitsainen, Irma and Päivi Pahta. 2003. "English in Finland: globalisation, language awareness and questions of identity". *English Today*; Cambridge 19: no. 4: 3-15. DOI:10.1017/S0266078403004024. Accessed 10 September 2020.
- Tajfel, H., and Turner, J. C. 1979. "An Integrative Theory of Intergroup Conflict". *The Social Psychology of Intergroup Relations*, eds. W. G. Austin and S. Worchel, 33–48. Monterey, CA: Brooks/Cole. Accessed 18 April 2020.
- Wearing, Michael. 2011. *Social Identity*. Nova Science Publishers Incorporated. ProQuest Ebook Central <http://ebookcentral.proquest.com/lib/kutu/detail.action?docID=3022005>. Accessed 18 April 2020.
- Yip, Virginia, and Stephen Matthews. 2007. *The Bilingual Child: Early Development and Language Contact*. Cambridge: Cambridge University Press. URL: [https://www-fulcrum-org.ezproxy.utu.fi/epubs/rx913q40h?locale=en#/6/98\[xhtml00000049\]!/4/4/1:0](https://www-fulcrum-org.ezproxy.utu.fi/epubs/rx913q40h?locale=en#/6/98[xhtml00000049]!/4/4/1:0). Accessed 4 September 2020.

Appendix 1.

Questionnaire for IB-students

1. Grade

pre-IB/1st/2nd

2. My mother tongue is

Finnish / English / Swedish / Other _____

3. My strongest language is

Finnish /English / Swedish / Other _____

4. My second strongest language is

Finnish / English / Swedish / Other _____

5. My third strongest language is

Finnish /English /Swedish /Other _____

6. The language used at home with other family members is

Finnish / Swedish / English / Other _____

7. I have attended English speaking (you can choose multiple options)

Day care / Elementary school / Middle school /High school / IB-program

8. I started learning English at the age of

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

9. I have lived in an English-speaking country (e.g. UK, US, Canada, Australia, New Zealand, South Africa) for

0-3 months

4-6 months

7-12 months

1-1.5 years

2+ years

None of the above

10. I have lived in a country other than Finland or an English-speaking country for

- 0-3 months
- 4-6 months
- 7-12 months
- 1-1.5 years
- 2+ years
- None of the above

Bilingualism (multilingualism) refers to the coexistence of more than one language system within an individual, as contrasted to monolingualism; and the regular use of two or more languages (or dialects) [mono- = one language, bi- = two languages, multi- = three or more]

11. Please select the most suitable option. 1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree

1. In my opinion, studying in English has been useful for me in order to define myself as a person.

1 2 3 4

2. In my opinion, I have reached the highest competence level (reading, writing, speaking and listening) I will be able to achieve in English.

1 2 3 4

3. In my opinion, I have made progress in learning English during my IB-studies

1 2 3 4

4. At school, I usually communicate with others in English also outside classroom

1 2 3 4

5. In my opinion, my friends consider me as bilingual/multilingual.

1 2 3 4

6. In my opinion, my family considers me as bilingual/multilingual.

1 2 3 4

12. Please select the most suitable option. 1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree

1. I feel more confident in English than in other languages.

1 2 3 4

2. I feel like a different person when I speak English.

1 2 3 4

3. I feel that my identity as an English speaker has become stronger during my IB-studies.

1 2 3 4

4. I feel like my identity as an English speaker has become weaker during my IB-studies.

1 2 3 4

5. I feel like my identity as a Finnish speaker has become stronger during my IB-studies.

1 2 3 4

6. I feel like my identity as a Finnish speaker has become weaker during my IB-studies.

1 2 3 4

13. Please select the most suitable option. 1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree

1. I feel bilingual when attending classes at school.

1 2 3 4

2. I feel bilingual when I am interacting with my family.

1 2 3 4

3. I feel bilingual when I am interacting with my friends.

1 2 3 4

4. I feel bilingual when I am travelling in English speaking countries.

1 2 3 4

5. I feel bilingual when I am travelling in non-English speaking countries.

1 2 3 4

6. I feel bilingual when I am speaking English with a native English speaker.

1 2 3 4

7. I feel bilingual when I am speaking English with a non-native English speaker.

1 2 3 4

8. I feel more bilingual now than at the beginning of my studies in IB

1 2 3 4

14. Please select the most suitable option. 1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree

1. I identify myself with the English-speaking community of the world.

1 2 3 4

2. I identify myself with native speakers of English.

1 2 3 4

3. I identify myself with non-native speakers of English.

1 2 3 4

4. I identify myself with monolingual Finnish speakers.

1 2 3 4

5. I identify myself with bilingual English speakers.

1 2 3 4

6. I identify myself with multilingual speakers.

1 2 3 4

15. According to the definition of bilingualism and the answers on this questionnaire, I categorize myself as...

Monolingual / Bilingual / Multilingual

16. Please choose one of the two sentences and continue it in your own words 1. "In my opinion, I *am* a bilingual, because..." 2. "In my opinion, I *am not* a bilingual, because..."

Appendix 2.

Finnish summary

Tämän pro gradu -tutkielman aihe on suomalaisissa IB-lukioissa opiskelevien nuorten subjektiivinen näkemys heidän kieli-identiteetistään ja siihen liittyvistä tekijöistä, sekä opiskelijoiden omat määritelmät kaksikielisyydelle. IB-lukion vaikutuksiin ja IB-lukiolaisiin liittyviä tutkimuksia ei Suomessa juuri ole tehty. IB-lukion nouseva suosio saattaa johtua englannin kielen yhä korkeammasta statuksesta suomalaisessa yhteiskunnassa, sillä englannin merkitys jokapäiväisen arjen jokaisella osa-alueella on kasvanut viime vuosikymmenien aikana huomattavasti (Leppänen & Nikula 2009). Jo vuosituhannen alussa on ennustettu, että Englanti tulee lähivuosina ohittamaan ruotsin tärkeysjärjestyksessä ja saavuttamaan Englanti toisena kielenä-statuksen (Taavitsainen and Pahta 2003, 1). Englannin merkityksen kasvu elämän jokaisella osa-alueella on saattanut vaikuttaa IB-lukion suosioon, sillä IB-lukio noudattaa kansainvälistä opetussuunnitelmaa, joka avaa ovia jatko-opiskeluun ulkomailla.

IB-lukiokoulutus on osa International Baccalaureate -järjestön tarjoamaa koulutusjärjestelmää, jonka opetussuunnitelmia auktorisoidut koulut saavat noudattaa ympäri maailman. Sen perusideana on tarjota CLIL-opetusta (eng. Content and Language Integrated Learning), mikä tarkoittaa, että opetus tapahtuu vieraalla kielellä. Suomessa vieraana kielenä toimii Englanti, joten IB-lukiolaiset käyttävät päivittäin Englantia opiskeluun. Tämän tutkimuksen tarkoituksena on selvittää, millainen IB-lukiolaisten kieli-identiteetti on ja miten taustatekijät, kuten kotona käytettävät kielet, ulkomailla asuminen tai luokka-aste vaikuttavat kieli-identiteettiin. Opiskelijoille annettiin mahdollisuus määrittellä oma kieli-identiteettinsä valitsemalla yksi kolmesta kieli-identiteettiryhmästä: yksikielinen, kaksikielinen ja monikielinen. Tämän avulla osallistujat jaettiin ryhmiin, ja näiden ryhmien välisiä eroja tarkasteltiin sen mukaan, miten ryhmien jäsenet määrittelivät kaksikielisyyden, ja millaisia eroja ryhmien välillä löytyi oman kielitaidon arvioinnissa, subjektiivisessa kaksikielisyyden kokemuksessa, kieliyhteisöihin identifioitumisessa ja identiteetin omaksumisessa. Näiden aiheiden perusteella muodostettiin kolme tutkimuskysymystä:

- 1) Millainen IB-lukiolaisten kieli-identiteetti yleisesti on?
- 2) Missä määrin taustatekijät vaikuttavat opiskelijoiden kieli-identiteettiin?
- 3) Missä määrin luokka-asteiden välillä on eroja kieli-identiteettiin liittyen?

Tutkielman hypoteesina on oletus, että suurin osa opiskelijoista määritteli itsensä yksikieliseksi, mikäli kielitaitoa ei koeta riittävän korkeaksi. Lisäksi yhteys englanninkielisiin

yhteisöihin ja aiempi englanninkielinen opetus saattaa vaikuttaa opiskelijoiden kieli-identiteettimääritelmiin. Koska identiteetti on muuttuva rakenne, kolmannen vuoden opiskelijat todennäköisemmin kategorisoivat itsensä kaksi- tai monikieliseksi, koska ovat opiskelleet englanniksi kauemmin kuin pre-IB-luokan opiskelijat, ja näin heidän kielitaitonsa ja identiteettinsä on saattanut vahvistua opiskeluiden myötä.

Tutkielman teoriaosiossa käsitellään identiteetin määritelmiä, muodostamisprosessia ja esitellään kieli-identiteetti osana identiteettiä. Identiteetti nähdään kielen oppimisen tutkimuksessa alati muuttavana kokonaisuutena, joka muodostetaan itsen ja sosiaalisen ympäristön välisessä vuorovaikutuksessa (Kanno 2003). Identiteetin vaikutus kielenoppimiseen on myös havaittu (Erikson 1956; Norton 2000). Koska identiteetin muodostaminen tapahtuu nuoruudessa ja varhaisaikuisuudessa, tämän tutkimuksen osallistujat saattavat vielä muodostaa identiteettiään (Erikson 1956; Marcia 1966). Identiteetti muodostuu henkilökohtaisesta identiteetistä ja sosiaalisesta identiteetistä. Henkilökohtainen identiteetti kuvaa sitä, mikä tekee ihmisestä uniikin yksilön ja muista eroavan (Saastamoinen 2006, 172). Sosiaalinen identiteetti muodostuu ryhmistä ja sosiaalisista kategorioista, joihin yksilö kokee kuuluvansa (Edwards 2008, 27). Identiteetti voi olla vakaa ja pysyvä, lyhytaikainen ja muuttuva tai tilannesidonnainen (Korostelina 2007, 61).

Sosiaaliseen identiteettiin kuuluu kieli-identiteetti, kulttuuri-identiteetti ja muut ryhmät, joihin yksilö kuuluu. Sosiaalinen identiteetti muodostetaan Tajfel ja Turnerin (1979) Sosiaalisen Identiteetin Teorian mukaan kolmen prosessin avulla: itsen ja ympäristön kategorisoiminen, samaistuminen ryhmän jäsenten kanssa ja näin identiteetin muodostaminen, ja lopulta oman sisäryhmän vertailu ulkoryhmiin. Kieli-identiteetti muodostetaan yksilön ja kielen välisessä vuorovaikutuksessa (Norton 2000). Näin ollen kieli-identiteetti on samaistumista kieliryhmän jäsenten kanssa. Tässä tutkielmassa kieli-identiteetti jaetaan kolmeen ryhmään: yksikieliseen, kaksikieliseen ja monikieliseen identiteettiin. Koska jaottelu on luonteeltaan liian kaksijakoinen, ja identiteetti ei välttämättä sovi valmiisiin kategorioihin, Kanno (2003) on yrittänyt ratkaista ongelmaa tarjoamalla uuden käsitteen: hybridi-identiteetti. Hybridi-identiteetti mahdollistaa identiteetin, joka on sekoitus useasta sisäryhmästä.

Kaksikielisyydelle on useita määritelmiä, mutta tässä tutkielmassa kaksikielisyys tarkoittaa kahden tai useamman kielen säännöllistä käyttöä, ja kaksikielinen on yksilö, joka jokapäiväisessä arjessaan käyttää kahta tai useampaa kieltä (Grosjean 2008). Yksikielisyys taas tarkoittaa vain yhden kielen käyttöä, ja monikielisyys useamman kuin kahden kielen käyttöä. Kaksikielisyysvaikutukset monet sosiaaliset tekijät, kuten kielten statukset yhteiskunnassa (Bhatia & Ritchie 2012,6), kielten erilaiset osa-alueet päivittäisessä elämässä ja lukutaito eri

kielissä (Hakuta 2010, 175) ja maahanmuutto (Bhatia & Ritchie 2012, 6). Yksilötasolla kaksikielisyyteen vaikuttavat ikä, jolloin kielille altistutaan (Hakuta 2010, 176-177; Yip & Matthews 2007, 25), kielitaito, (Bhatia & Ritchie 2012, 12) ja aktiivinen vs. passiivinen kaksikielisyys (Hakuta 2010, 177). Lisäksi Grosjean (2008) on jakanut kaksikielisyyden neljään kategoriaan, joilla pyritään määrittelemään, miten kaksikieliset nähdään suhteessa kielenkäyttöön ja kulttuuriin. Kategoriat ovat yksikielinen käsitys kaksikielisyydestä (kaksikielinen on kaksi yksikielistä yhdessä yksilössä), kaksikielinen käsitys kaksikielisyydestä (kaksikielisen kielisysteemit ovat yksi iso kokonaisuus), yksikulttuurinen kaksikielisyys (kaksikielinen samaistuu vain toiseen kulttuuriin) ja kaksikulttuurinen kaksikielisyys (kaksikielinen samaistuu molempiin kulttuureihin). Näiden avulla pyritään määrittelemään se taso, jolla yksilö samaistuu kieliryhmiin ja kulttuureihin.

Kieli-identiteettiin ja kaksikielisyyteen liittyvien teorioiden avulla on muodostettu kategoriat, joilla kaksikielisyyttä voidaan kuvata. Kaksikielisyyden määritelmiin kuuluvat siis ikä, kielitaso (syntyperäisyys, sujuvuus, eri osa-alueiden taitotasot), kielenkäytön säännöllisyys ja jaottelu eri elämän osa-alueille, affektiiviset tekijät kuten tunteet, samaistuminen kieliryhmiin ja kulttuureihin, ja monikielisyys. Näihin kategorioihin perustuu sekä kysely että avokysymyksen vastausten sisältöanalyysi.

Tutkielman teoriataustan toinen osa käsittelee CLIL-opetusta ja sen vaikutuksia, sekä esittelee IB-lukion CLIL-opetusta noudattavana koulutuksena. CLIL (eng. Content and Language Integrated Learning) on yläkäsite, joka tarkoittaa opetusta, jossa käytetään pääasiassa vierasta kieltä (Nikula 2007, 208-209). CLIL-opetuksessa kieli ei enää ole oppimisen kohde, vaan oppimisen ja kommunikaation mahdollistaja (Halbach 2014,2). CLIL-opetuksen vaikutuksia on tutkittu runsaasti muualla Euroopassa, ja tutkitusti CLIL-opetus parantaa motivaatiota kielenoppimiseen (emt.) ja kehittää kommunikaatiostrategioita (Coyle 2007, 553), kehittää kielitaitoa (Pérez, Luisa and Lancaster 2017), ja usein saattaa jopa parantaa sisältöjen oppimista (Surmont et al. 2016). Lisäksi CLIL-opetuksella on havaittu olevan vaikutuksia turhautumisensietokykyyn ja itsevarmuuteen (Vollmer et al.in Salazar-Noguera 2016, 55). CLIL-opetuksen vaikutukset saattavat kuitenkin riippua yksilöstä.

IB-organisaatio on maailmanlaajuinen organisaatio, joka tarjoaa kansainvälisiä koulutusohjelmia neljällä eri tasolla. Näistä Diploma Programme on ohjelma, jota tarjotaan Suomessa kansallisen lukion rinnalla (ibo.org). Suomen IB-lukiot noudattavat organisaation opetussuunnitelmaa, jossa tutkinto suoritetaan vieraalla kielellä. Näin IB-lukiot noudattavat CLIL-opetusmetodia. IB-lukion tarjoamat jatko-opintomahdollisuudet ovat laajemmat, sillä IB-tutkinto on suoraan vertailukelpoinen ulkomaisissa yliopistoissa (ibo.org). IB-lukio on

kasvattanut suosiotaan, ja 30 vuoden toiminnan jälkeen IB-lukiolaisia on Suomessa yli 1000 (syk.fi).

CLIL-opetuksen vaikutuksia kieli-identiteettiin on tutkittu hyvin vähän. Koska kieli-identiteetti muodostetaan yksilön ja sosiaalisen ympäristön välisessä vuorovaikutuksessa ja se tehdään kielen avulla, kieli on suuressa osassa identiteetin muodostamisprosessissa (Kanno 2003). Yksilö siis käyttää kieltä identiteetin muodostamiseen, ja näin antaa yksilölle valtaa valita ja muokata identiteettiään (emt.) Uusien kielten oppiminen saattaa siis muuttaa yksilön kieli-identiteettiä antamalla valtaa ja mahdollistamalla uusien identiteettien muodostamisen (Grosjean 2008). Lisäksi, koska kielen oppimisen sosiaalinen ja kulttuurinen konteksti saattaa vaikuttaa kieli-identiteettiin, IB-lukiossa opiskelemisella saattaa olla vaikutuksia opiskelijoiden identiteettiin tarjoamalla kontekstin uudenlaisen identiteetin muodostamiselle (Higgins 2011). Aiemmat tutkimukset CLIL-opetuksen vaikutuksista Suomessa ovat olleet kvalitatiivisia. Esimerkiksi McCambridge (2007) ja Kovanen (2011) tutkivat tutkielmissaan IB-lukiolaisia haastatteleamalla heitä. Molempien tutkimuksissa kävi ilmi, että tutkittavat eivät pitäneet itseään kaksikielisinä, vaikka kokivat englannin olevan suuri osa elämää.

Tutkielman tiedonkeruumenetelmänä käytettiin Webropol-kyselyä, joka koostui monivalintakysymyksistä taustatietoihin liittyen, neliasteisista Likert-asteikkokysymyksistä kaksikielisyyden subjektiivisiin kokemuksiin liittyen, ja avokysymyksestä, jossa osallistujia pyydettiin kertomaan, miksi kokee tai ei koe olevansa kaksikielinen. Kysely muodostettiin Kasurisen (2019) kyselyn ja Dörnyein ja Taguchin (2009) teoksen avulla. Kyselylinkki lähetettiin sähköpostitse suomalaisiin IB-lukioihin, joista kolmesta koulusta vastattiin. Vastauksia saatiin yhteensä 114.

Tulokset analysoitiin sekä kvantitatiivisesti käyttämällä IBM SPSS Statistics -ohjelmistoa että kvalitatiivisesti sisältöanalyysia hyväksikäyttäen. Kvantitatiivisessa analyysissa vertailtiin eroja subjektiivisissa kokemuksissa kieli-identiteettiryhmien välillä ja vertailtiin taustatekijöiden ja luokka-asteen vaikutuksia opiskelijoiden kieli-identiteettiin. Sisältöanalyysi keskittyi vastaajien omiin määritelmiin kieli-identiteetistään, ja jaottelu perustui teorian pohjalta muodostettuihin kaksikielisyyden kategorioihin. Vastaajat jaettiin ryhmiin oman kategorisointinsa perusteella: yksikielisiin, kaksikielisiin ja monikielisiin. Lisäksi yksikielisten ja ei-yksikielisten (=kaksikieliset ja monikieliset) välisiä eroja analysoitiin.

Tulokset esitettiin teemoittain. Tulokset paljastivat, että vain 4% vastaajista (N=5) kategorisoi itsensä yksikieliseksi, kun taas kaksikieliseksi itsensä määrittelivät 55% (N=63) ja monikieliseksi 41% (N=46) vastaajista. Tulokset eriävät aiemmista tutkimuksista, joissa osallistujat mielsivät itsensä yksikieliseksi (McCambridge 2007; Kovanen 2011). Vastaajista 16

raportoi puhuvansa kahta kieltä äidinkielenään, ja tulokset paljastivat kielirepertuaareissa samankaltaisuuksia: yleisin kieliyhdistelmä IB-lukiolaisten keskuudessa on suomi-englanti-ruotsi. Eniten vaihtelua kieliyhdistelmissä löytyi monikielisten ryhmästä, ja ruotsin asema näyttäytyi kieliyhdistelmissä verrattain heikkona kaikissa ryhmissä (vahvin kieli 12%:lla, toiseksi vahvin 9%:lla ja kolmanneksi vahvin 42%:lla kaikista vastaajista). Tulokset vahvistavat näkemystä siitä, että englantia on saavuttamassa tai jo saavuttanut osassa suomalaista populaatiota korkeamman aseman kuin virallinen toinen kieli ruotsi (Taavitsainen ja Pahta 2003). Teemaan liittyen, osattujen kielten määrä (N=13) ja kielen omaksumisikä tai äidinkielen määrä (N=9) olivat usein mainittuja kaksikielisyyden määritelmiä. Varsinkin itsensä monikieliseksi kategorisoineet osallistujat perustelivat kaksikielisyyttään kielten lukumäärällä (N= 10).

Tulososion toisessa osiossa analysoitiin kaksikielisyyden subjektiivisia kokemuksia Likert-asteikkokysymysten vastauksiin perustuen. Analysoinnissa käytettiin tilastollisia Kruskal-Wallis -ja Mann-Whitney U-testejä, jotta vertailu ryhmien välillä oli mahdollista. Tilastollisesti merkittäviä eroja löydettiin. Tulosten perusteella kaksi- ja monikieliset ajattelevat myös läheistensä pitävän heitä kaksikielisinä, kun taas yksikieliset ajattelevat päinvastoin. Lisäksi kaksi- ja monikielisistä tuntuu erilaiselta englantia puhuessa, he tuntevat itsensä vahvemmin kaksikieliseksi kommunikoidessaan muiden kanssa ja matkustaessaan ulkomailla, ja identifioivat itsensä vahvimmin omaa identiteettiä vastaavan puhujaryhmän kanssa (kaksikieliset kaksikielisten kanssa, monikieliset monikielisten kanssa). Lisäksi, vaikka tilastollista merkitsevyyttä ei löytynyt, kaksi- ja monikieliset kokivat identiteettinsä muuttuneen IB-opintojen aikana: englannin puhujan identiteetistä on tullut vahvempi, kun taas suomen puhujan identiteetti on heikentynyt. Tulokset ovat linjassa teoriaosiossa esiteltyjen teorioiden kanssa identiteetistä ja kaksikielisyydestä (esim. Korostelina 2007; Grosjean 2008; Hakuta 2010). Yksikieliset taas kokivat, että heillä oli enemmän kehitettävää kielitaitonsa suhteen, eivätkä kokeneet kieli-identiteetissään muutosta IB-opiskelujen aikana. Lisäksi yksikieliset eivät identifioineet itseään yhdenkään tarjotun kieliryhmän kanssa, mikä saattaa olla hybridi-identiteetin vaikutusta (Kanno 2003). Lisäksi, tulokset osoittavat, että mikään kolmesta ryhmästä ei tunne itseään varmaksi englannin käyttäjäksi. Subjektiivisista kaksikielisyyden määritelmistä yleisimmin käytettiin kielitaitoon liittyviä perusteluja (kielitaito N=43, sujuvuus N=38, syntyperäisyys N=6) ja kielenkäytön frekvenssiä (N=29). Affektiiviset tekijät mainittiin neljä kertaa, koodinvaihto kolme kertaa ja itsevarmuus 3 kertaa aineistossa.

Taustatekijöiden vaikutuksista tilastollinen merkitsevyys löytyi kieli-identiteetin ja kotona käytettyjen kielten lukumäärän väliltä. Tulosten perusteella kotona käytettyjen kielten

lukumäärä vaikuttaa valintaan kaksikielisen ja monikielisen identiteetin välillä. Muiden taustatekijöiden tai luokka-asteen ja kieli-identiteetin välillä ei löytynyt merkitsevyyttä.

Tulosten perusteella voidaan todeta, että suomalaiset IB-opiskelijat määrittelevät itsensä yleisemmin joko kaksi- tai monikieliseksi. Kieli-identiteettiryhmien välillä löydettiin merkittäviä eroja muiden näkemyksistä omasta kaksikielisyydestä, kaksikielisyyden subjektiivisissa näkemyksissä, identiteetin muodostamisessa ja kieliryhmiin samaistumisessa. Lisäksi ruotsin aseman heikkeneminen toisesta kielestä kolmanteen kieleen havaittiin tuloksissa. Sisältöanalyysin perusteella havaittiin, että yleisimmin opiskelijat määrittelivät kaksikielisyyttään kielitaidolla ja kielten lukumäärällä sekä kielenkäytön frekvenssillä, jotka ovat linjassa kaksikielisyyden määritelmien kanssa.

Tutkielma avaa kieli-identiteetin käsitettä monimutkaisena sosiaalisena ilmiönä, johon vaikuttavat monet yksilötason tekijät, mutta myös sosiaaliset tekijät ovat suuressa osassa identiteetin muodostamisessa. Lisäksi tutkielma antaa osviittaa siitä, mihin suuntaan suomalaisten IB-lukiolaisten kieli-identiteetti on kehittymässä ja tarjoaa ideoita jatkotutkimukselle. Suuremmalla ja tasapainoisemmin jakautuneella otoksella voitaisiin tutkia paremmin taustatekijöiden vaikutusta kieli-identiteettiin. Lisäksi, koska yksikieliset eivät samaistuneet mihinkään kieliryhmään, yksikielisten identiteetti voisi olla mielenkiintoinen tutkimuskohde. Toisaalta, koska identiteetti on muuttuva rakenne, pitkittäistutkimus kieli-identiteetin kehityksestä saattaisi antaa arvokasta tietoa kieli-identiteetistä.